10088269 .031602

GLM:dsh 03/15/02 4239-62295 E-272-99/2

EXPRESS MAIL LABEL NO. EV053212979US
DATE OF DEPOSIT: March 15, 2002

JC10 Rec'd PCT/PTO 15 MAR 2002

FORM PTO-1390 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. § 371		4239-62295 U.S. APPLICATION NO. (If known, see 37 C.F.R. § 1.5) Unknown 10/088269
INTERNATIONAL APPLICATION NO. PCT/US00/25465	INTERNATIONAL FILING DATE September 15, 2000	PRIORITY DATE CLAIMED September 17, 1999
TITLE OF INVENTION SIGNAL COUNTING FOR IN SITU HYBRIDIZATION		
APPLICANT(S) FOR DO/EO/US Olli P. Kallioniemi, Juha Kononen, Lukas Bubendorf, Edward R. Dougherty, and Artyom M. Grigoryan		

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. § 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. § 371.
3. ☒ This is an express request to begin national examination procedures (35 U.S.C. § 371(f) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. § 371(b) and PCT Articles 22 and 39(1).
4. ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. § 371(c)(2))
   a. ☒ is transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☐ has been transmitted by the International Bureau.
   c. ☒ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ A translation of the International Application into English (35 U.S.C. § 371(c)(2)).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. § 371(c)(3))
   a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☐ have been transmitted by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☒ have not been made and will not be made.
8. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. § 371(c)(3)).
9. ☒ An oath or declaration of the inventor(s) (35 U.S.C. § 371(c)(4)).
10. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. § 371(c)(5)).

Items 11. to 16. below concern document(s) or information included:

11. ☐ An Information Disclosure Statement under 37 C.F.R. §§ 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 C.F.R. §§ 3.28 and 3.31 and the Recordal fee of $40.00 is included.
13. ☒ A **FIRST** preliminary amendment.
    ☐ A **SECOND** or **SUBSEQUENT** preliminary amendment.
14. ☐ A substitute specification.
15. ☐ A change of power of attorney and/or address letter.
16. ☒ Other items or information:
    ☒ Abstract on a separate page.
    ☐ Written Opinion.
    ☐ Preliminary Examination Report.
    ☐ International Search Report.
    ☐ Copies of References Cited.

24197

10088269 .031602

GLM:dsh 03/15/02 4239-62295 E-272-99/2

EXPRESS MAIL LABEL NO. EV053212979US
DATE OF DEPOSIT: March 15, 2002

JC10 Rec'd PCT/PTO 15 MAR 2002

U.S. APPLICATION NO. (If known, see 37 C.F.R. 1.5)	INTERNATIONAL APPLICATION NO.	ATTORNEY'S DOCKET NUMBER
Unknown 10/088269	PCT/US00/25465	4239-62295

17. ☒ The following fees are submitted:

CALCULATIONS (PTO USE ONLY)

BASIC NATIONAL FEE (37 C.F.R. §§ 1.492(a)(1)-(5)):

Neither International Preliminary Examination fee (37 C.F.R. § 1.482) nor International Search fee (37 C.F.R. § 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO **$1,040.00**

International Preliminary Examination fee (37 C.F.R. § 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO................................ **$890.00**

International Preliminary Examination fee (37 C.F.R. § 1.482) not paid to USPTO but International Search fee (37 C.F.R. § 1.445(a)(2)) paid to USPTO as an International Searching Authority **$740.00**

International Preliminary Examination fee paid to USPTO (37 C.F.R. § 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4)...................................... **$710.00**

International Preliminary Examination fee paid to USPTO (37 C.F.R. § 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4) **$100.00**

ENTER APPROPRIATE BASIC FEE AMOUNT =				**$ 890.00**	
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☒ 30 months from the earliest claimed priority date (37 C.F.R. § 1.492(e)).				**$ 130.00**	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total claims	46 - 20 =	26	**x $18.00**	**$ 468.00**	
Independent Claims	3 - 3 =	0	**x $84.00**	**$ 0**	
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			**+ $280.00**	**$**	
TOTAL OF ABOVE CALCULATIONS =				**$ 1,488.00**	
☐ Reduction of 1/2 for filing by small entity. Small entity status is claimed for this application.				**$**	
SUBTOTAL =				**$ 1,488.00**	
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 Months from the earliest claimed priority date (37 C.F.R. §§ 1.492(f)). +				**$**	
TOTAL NATIONAL FEE =				**$ 1,488.00**	
Fee for recording the enclosed assignment (37 C.F.R. § 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 C.F.R. §§ 3.28, 3.31). **$40.00** per property. +				**$**	
TOTAL FEES ENCLOSED =				**$ 1,488.00**	
				REFUND →	**$**
				CHARGE →	**$**

a. ☒ A check in the amount of $ 1,488.00 to cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. _____ in the amount of $ _____ to cover the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Director is hereby authorized to charge any additional fees that may be required, or credit any overpayment, to Deposit Account No. 02-4550. A duplicate copy of this sheet is enclosed.

d. ☒ Please return the enclosed postcard to confirm that the items listed above have been received.

NOTE: **Where an appropriate time limit under 37 C.F.R. § 1.494 or § 1.495 has not been met, a petition to revive (37 C.F.R. § 1.137(a) or (b)) must be filed and granted to restore the application to pending status.**

SEND ALL CORRESPONDENCE TO:

KLARQUIST SPARKMAN, LLP
One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, OR 97204-2988

G. L. Maurer
SIGNATURE
Gregory L. Maurer
NAME

43,781
REGISTRATION NUMBER

cc: Docketing

10/088269

GLM:dsh 03/15/02 E-272-99/1
PATENT

Attorney Reference Number 4239-62295
Express Mail No. EV053212979

JC10 Rec'd PCT/PTO 15 MAR 2002

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Kallioniemi et al

Application No. Not yet assigned

Filed: Herewith

For: SIGNAL COUNTING FOR IN SITU HYBRIDIZATION

Examiner: Not yet assigned

Date: March 15, 2002

Art Unit: Not yet assigned

CERTIFICATE OF MAILING

I hereby certify that this paper and the documents referred to as being attached or enclosed herewith are being deposited with the United States Postal Service on March 15, 2002 as Express Mail No. EV053212979 in an envelope addressed to: BOX PCT COMMISSIONER FOR PATENTS, WASHINGTON, D.C. 20231.

Attorney for Applicant

BOX PCT
COMMISSIONER FOR PATENTS
WASHINGTON, D.C. 20231

PRELIMINARY AMENDMENT

Prior to calculating the claim fees for the above-identified patent application, please amend the application as follows to comply with national stage requirements and otherwise amend the claims. Please disregard and do not enter the Article 34 Amendment filed during international prosecution, which the Applicants do not wish entered at this time.

In the Specification:

On page 1, after the title, please insert the following paragraph:

--Priority Claim

This is a § 371 U.S. national stage of PCT/US00/25465, filed September 15, 2000, which was published in English under PCT Article 21(2), and claims the benefit of U.S. Application No. 60/154,601, filed September 17, 1999.--

Please place the following abstract (which is also submitted on a separate page, attached) at the end of the specification (i.e., as page 45):

--SIGNAL COUNTING FOR IN SITU HYBRIDIZATION

ABSTRACT

A computer system counts fluorescently tagged nucleic acid probe signals in biological specimens by determining a ratio of signals from a test probe to signals of a reference probe. Probe signals need not be counted with reference to cells, nuclei, or nuclear contours. Gene amplification or deletion can thus be detected by analyzing the ratio. Successive image slices are obtained by confocal microscopy, and the images are digitized. The digital images are transformed and analyzed to combine contiguous fluorescent signal segments in successive optical sections to identify discrete probe signals, or spots. Spots overlapping in the axial and transverse dimensions of a three-dimensional representation of the biological specimens can be distinguished. A graphical user interface presents various features for consideration by a user, who can provide guidance to a computer system counting the spots. Various features directed to identifying spot clusters and autofluorescent material can increase accuracy of spot counting.--

In the Claims:

Please cancel claims 41-57 without prejudice.

Please amend claims 1, 15, 24 as follows:

1. (Amended) A computer-implemented method for counting nucleic acid probe signals in a region of interest in a biological specimen, the method comprising:

in a computer system, automatically counting a number of test signals from a test probe;

in the computer system, automatically counting a number of reference signals from a reference probe; and

in the computer system, determining a ratio of the automatically-counted test signals from the test probe to the automatically-counted reference signals from the reference probe, wherein the region of interest comprises multiple cells.

2. The method of claim 1, wherein the reference probe is a polynucleotide that hybridizes to a centromere, and the number of reference signals from the reference probe approximates a nucleus count in the biological specimen.

3. The method of claim 1, wherein the reference probe recognizes a target on a same chromosome as the test probe.

4. The method of claim 1, wherein the test probe is a polynucleotide that hybridizes to a target sequence in a gene, and the reference probe is a polynucleotide that hybridizes to a reference sequence.

5. The method of claim 3, wherein the reference probe recognizes a centromere of the same chromosome on which the gene of interest is contained.

6. The method of claim 1, further comprising obtaining successive images of the region of interest to distinguish overlapping signals in the biological specimen.

7. The method of claim 6, wherein the successive images are optical sections of the region of interest.

8. The method of claim 7, wherein the optical sections are at different depths of the biological specimen.

9. The method of claim 8, wherein the successive images are transformed into digital representations in which contiguous signal segments in successive optical sections are combined into a single signal in a particular optical section in which a strongest signal segment is located.

10. The method of claim 6, wherein different successive images are obtained for the test probe signals and the reference probe signals, and a quantity of test probe signals and reference probe signals are determined.

11. The method of claim 6, wherein successive images are obtained which show distinguishable test probe signals and reference probe signals, and a quantity of the test probe signals and reference probe signals are determined.

12. The method of claim 6, wherein the successive images are obtained by confocal microscopy.

13. The method of claim 1, wherein the ratio of signals is determined without reference to boundaries of a cell nucleus.

14. The method of claim 1, wherein the ratio of signals is determined without reference to the boundaries of a cell.

15. (Amended) The method of claim 1 wherein the probe signals are visible signals from probes used with in situ hybridization of a biological sample, the method further comprising:

obtaining a plurality of images at different levels of the biological sample; and

constructing a three-dimensional image indicating discrete signals at different levels of the three-dimensional image;

wherein automatically counting comprises counting computer-identified discrete signals out of the discrete signals at different levels of the three-dimensional image.

16. The method of claim 15, wherein the three-dimensional image is constructed by determining a location of a signal segment in the different levels of the biological sample, combining overlapping signal segments in contiguous levels into a single spot signal, and separating signal segments in non-contiguous levels into different spots.

17. The method of claim 16, wherein the location of signal segments is determined by the presence of an increase in brightness intensity that indicates an increase of signal as compared to a background signal.

18. The method of claim 17, wherein the probes display fluorescent signals, and the increase in brightness intensity is associated with an increase in fluorescence compared to the background signal.

19. The method of claim 15, wherein the signals comprise test signals from a test probe and reference signals from a reference probe.

20. The method of claim 19, wherein the test probe recognizes a gene of interest, and the reference probe recognizes a chromosomal locus having an expected quantity in the biological specimen.

21. The method of claim 20, further comprising determining a ratio between the test signals and the reference signals.

22. The method of claim 21, further comprising determining:

(a) whether there is an increase in an expected ratio between the test signal and the reference signal, indicating an amplification of the gene of interest; or

(b) whether there is a decrease in the expected ratio between the test signal and the reference signal, indicating relative loss of the gene of interest.

23. The method of claim 19, wherein the test probe is selected from the group consisting of probes that recognize genes implicated or suspected in the development or progression of a tumor.

24. (Amended) The method of claim 15, wherein the biological sample is in a microarray.

25. The method of claim 24, wherein the microarray comprises a tissue microarray.

26. The method of claim 25, wherein the tissue microarray comprises tissue samples of a same tissue type taken from a plurality of donor specimens.

27. The method of claim 15, wherein the plurality of images consists of between eight and thirty two images at different levels of the biological sample.

28. The method of claim 15, further comprising:

avoiding counting discrete signals having intensities exceeding a threshold intensity.

29. The method of claim 15, further comprising:

avoiding counting discrete signals having a combined intensity and area exceeding a threshold value.

30. The method of claim 15, further comprising:

avoiding counting discrete signals related to autofluorescent material.

31. The method of claim 15, further comprising:

depicting a two-dimensional image representing the three-dimensional image for consideration by a user.

32. The method of claim 31, further comprising:

emphasizing discrete signals related to autofluorescent material in the two-dimensional image.

33. The method of claim 15, further comprising:

identifying a set of one or more discrete signals as a cluster; and

counting the cluster as a number of discrete signals greater than the number of discrete signals in the set.

34. The method of claim 33 wherein the cluster is counted as a number of discrete signals indicated by applying a mapping to the number of discrete signals in the set.

35. The method of claim 33 wherein the cluster is counted as a number of discrete signals indicated by a function calibrated via manual counting of spots in a plurality of images.

36. The method of claim 33 wherein the cluster is counted as a number of discrete signals indicated by a gain factor applied to the number of discrete signals in the set.

37. The method of claim 15 wherein the plurality of images are a set of images taken during a first analysis of a first color channel, and a second set of images are taken of the biological sample for a second color channel, the method further comprising:

avoiding counting discrete signals appearing at a same location in the set of images for the first color channel and the set of images in the second color channel.

38. The method of claim 15 wherein the plurality of images are a set of images taken for a test probe, and a second set of images are taken of the biological sample for a reference probe, the method further comprising:

avoiding counting discrete signals appearing at a same location in the set of images for the test probe and the set of images for the reference probe.

39. The method of claim 15 further comprising:

receiving a directive from a user indicating counting is to be avoided for a specified portion of the biological sample; and

responsive to the directive, avoiding counting discrete signals for the specified portion of the biological sample.

60. A computer-generated user interface for presenting results of microscopic observation of biological tissue subjected to a FISH experiment, the user interface comprising:

a scatter plot of sets of image components designated as spot candidates for the FISH experiment;

wherein the scatter plot comprises points indicating a size and intensity of spot candidates.

61. The computer-generated user interface of claim 60 wherein at least one point is operable to receive a user interface activation to navigate to a user interface display of information for a spot candidate associated with the point.

62. The computer-generated user interface of claim 60 wherein at least one point is operable to receive a user interface activation to navigate to a user interface display of a three-dimensional depiction of a spot candidate associated with the point.

63. The computer-generated user interface of claim 60 further comprising:

a display image depicting a view of the tissue subjected to the FISH experiment and a depiction of least one candidate spot thereon;

wherein the depiction of the candidate spot is operable to receive a user interface activation to designate the spot as a minimal intensity spot; and

wherein candidate spots designated as minimal intensity spots are visually emphasized when presenting the scatter plot.

REMARKS

The specification has been amended herein to insert Applicant's claim of priority, and to insert an Abstract as the last page of the specification. The claims have been amended for clarification. No new matter has been added.

The priority claim to U.S. Provisional Application No. 60/154,601 was already set forth in the cover sheet that accompanied the PCT patent application (Application No. PCT/US00/25465). The priority claim to the PCT is set forth in the cover sheet

40. The method of claim 15 further comprising:

receiving a directive from a user indicating counting is to be performed separately for a specified portion of the biological sample; and

responsive to the directive, separately counting discrete signals for the specified portion of the biological sample.

41-57 (Canceled)

58. A computer-readable medium comprising computer-executable instructions for performing the following:

within a stack of image slices generated from a plurality of confocal microscopic observations of a FISH experiment as a plurality of depths along a z-axis, identifying possible fluorescent image components;

projecting the possible fluorescent image components within the image slices onto a projection image;

discarding insignificant contiguous possible fluorescent image components in the slices;

for each contiguous region in the projection image, grouping regions of possible florescent image components associated with the contiguous region in the projection image into spot candidates;

applying a filter to the spot candidates; and

counting the remaining spot candidates as spots.

59. The computer-readable medium of claim 58 wherein insignificant contiguous possible fluorescent image components are determined by comparing a size of a contiguous possible fluorescent image component with a threshold.

accompanying the patent application filed herewith and in the copy of the declaration that is submitted with the application filed herewith.

If any matters remain to be discussed prior to examination, the Examiner is invited to contact the undersigned at the telephone number listed below.

Respectfully submitted,

KLARQUIST SPARKMAN CAMPBELL
LEIGH & WHINSTON, LLP

By ______________________
Gregory L. Maurer
Registration No. 43,781

One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, Oregon 97204
Telephone: (503) 226-7391
Facsimile: (503) 228-9446

Marked-up Version of Amended Specification and Claims
Pursuant to 37 C.F.R. §§ 1.121(b)-(c)

***In the Specification*:**

On page 1, after the title, the following paragraph has been inserted:

<u>Priority Claim</u>

<u>This is a § 371 U.S. national stage of PCT/US00/25465, filed September 15, 2000, which was published in English under PCT Article 21(2), and claims the benefit of U.S. Application No. 60/154,601, filed September 17, 1999.</u>

In the Abstract:

The attached abstract has been added to the end of the specification.

In the Claims:

The claims have been amended as follows:

1. **(Amended)** A **<u>computer-implemented</u>** method for counting nucleic acid probe signals in a region of interest in a biological specimen, the method comprising:

<u>in a computer system, automatically</u> counting a number of test signals from a test probe;

<u>in the computer system, automatically</u> counting a number of reference signals from a reference probe; and

<u>in the computer system,</u> determining a ratio of the **<u>automatically-</u>**counted test signals from the test probe to the **<u>automatically-</u>**counted reference signals from the reference probe, wherein the region of interest comprises multiple cells.

2. The method of claim 1, wherein the reference probe is a polynucleotide that hybridizes to a centromere, and the number of reference signals from the reference probe approximates a nucleus count in the biological specimen.

3. The method of claim 1, wherein the reference probe recognizes a target on a same chromosome as the test probe.

4. The method of claim 1, wherein the test probe is a polynucleotide that hybridizes to a target sequence in a gene, and the reference probe is a polynucleotide that hybridizes to a reference sequence.

5. The method of claim 3, wherein the reference probe recognizes a centromere of the same chromosome on which the gene of interest is contained.

6. The method of claim 1, further comprising obtaining successive images of the region of interest to distinguish overlapping signals in the biological specimen.

7. The method of claim 6, wherein the successive images are optical sections of the region of interest.

8. The method of claim 7, wherein the optical sections are at different depths of the biological specimen.

9. The method of claim 8, wherein the successive images are transformed into digital representations in which contiguous signal segments in successive optical sections are combined into a single signal in a particular optical section in which a strongest signal segment is located.

10. The method of claim 6, wherein different successive images are obtained for the test probe signals and the reference probe signals, and a quantity of test probe signals and reference probe signals are determined.

11. The method of claim 6, wherein successive images are obtained which show distinguishable test probe signals and reference probe signals, and a quantity of the test probe signals and reference probe signals are determined.

12. The method of claim 6, wherein the successive images are obtained by confocal microscopy.

13. The method of claim 1, wherein the ratio of signals is determined without reference to boundaries of a cell nucleus.

14. The method of claim 1, wherein the ratio of signals is determined without reference to the boundaries of a cell.

15. **(Amended)** [A method of counting] **The method of claim 1 wherein the probe signals are** visible signals from probes used with in situ hybridization of **a** biological [samples] **sample**, the method **further** comprising:

obtaining a plurality of images at different levels of the biological sample; **and**

constructing a three-dimensional image indicating discrete signals at different levels of the three-dimensional image; **~~and~~**

wherein automatically counting comprises counting **computer-identified discrete signals out of** the discrete signals at different levels of the three-dimensional image.

16. The method of claim 15, wherein the three-dimensional image is constructed by determining a location of a signal segment in the different levels of the biological sample, combining overlapping signal segments in contiguous levels into a single spot signal, and separating signal segments in non-contiguous levels into different spots.

17. The method of claim 16, wherein the location of signal segments is determined by the presence of an increase in brightness intensity that indicates an increase of signal as compared to a background signal.

18. The method of claim 17, wherein the probes display fluorescent signals, and the increase in brightness intensity is associated with an increase in fluorescence compared to the background signal.

19. The method of claim 15, wherein the signals comprise test signals from a test probe and reference signals from a reference probe.

20. The method of claim 19, wherein the test probe recognizes a gene of interest, and the reference probe recognizes a chromosomal locus having an expected quantity in the biological specimen.

21. The method of claim 20, further comprising determining a ratio between the test signals and the reference signals.

22. The method of claim 21, further comprising determining:

(a) whether there is an increase in an expected ratio between the test signal and the reference signal, indicating an amplification of the gene of interest; or

(b) whether there is a decrease in the expected ratio between the test signal and the reference signal, indicating relative loss of the gene of interest.

23. The method of claim 19, wherein the test probe is selected from the group consisting of probes that recognize genes implicated or suspected in the development or progression of a tumor.

24. **(Amended)** The method of claim 15, wherein the biological [samples are] **sample is** in a microarray.

25. The method of claim 24, wherein the microarray comprises a tissue microarray.

26. The method of claim 25, wherein the tissue microarray comprises tissue samples of a same tissue type taken from a plurality of donor specimens.

27. The method of claim 15, wherein the plurality of images consists of between eight and thirty two images at different levels of the biological sample.

28. The method of claim 15, further comprising:

avoiding counting discrete signals having intensities exceeding a threshold intensity.

29. The method of claim 15, further comprising:

avoiding counting discrete signals having a combined intensity and area exceeding a threshold value.

30. The method of claim 15, further comprising:

avoiding counting discrete signals related to autofluorescent material.

31. The method of claim 15, further comprising:

depicting a two-dimensional image representing the three-dimensional image for consideration by a user.

32. The method of claim 31, further comprising:

emphasizing discrete signals related to autofluorescent material in the two-dimensional image.

33. The method of claim 15, further comprising:

identifying a set of one or more discrete signals as a cluster; and

counting the cluster as a number of discrete signals greater than the number of discrete signals in the set.

34. The method of claim 33 wherein the cluster is counted as a number of discrete signals indicated by applying a mapping to the number of discrete signals in the set.

35. The method of claim 33 wherein the cluster is counted as a number of discrete signals indicated by a function calibrated via manual counting of spots in a plurality of images.

36. The method of claim 33 wherein the cluster is counted as a number of discrete signals indicated by a gain factor applied to the number of discrete signals in the set.

37. The method of claim 15 wherein the plurality of images are a set of images taken during a first analysis of a first color channel, and a second set of images are taken of the biological sample for a second color channel, the method further comprising:

avoiding counting discrete signals appearing at a same location in the set of images for the first color channel and the set of images in the second color channel.

38. The method of claim 15 wherein the plurality of images are a set of images taken for a test probe, and a second set of images are taken of the biological sample for a reference probe, the method further comprising:

avoiding counting discrete signals appearing at a same location in the set of images for the test probe and the set of images for the reference probe.

39. The method of claim 15 further comprising:

receiving a directive from a user indicating counting is to be avoided for a specified portion of the biological sample; and

responsive to the directive, avoiding counting discrete signals for the specified portion of the biological sample.

40. The method of claim 15 further comprising:

receiving a directive from a user indicating counting is to be performed separately for a specified portion of the biological sample; and

responsive to the directive, separately counting discrete signals for the specified portion of the biological sample.

41-57 **(Canceled)**

58. A computer-readable medium comprising computer-executable instructions for performing the following:

within a stack of image slices generated from a plurality of confocal microscopic observations of a FISH experiment as a plurality of depths along a z-axis, identifying possible fluorescent image components;

projecting the possible fluorescent image components within the image slices onto a projection image;

discarding insignificant contiguous possible fluorescent image components in the slices;

for each contiguous region in the projection image, grouping regions of possible florescent image components associated with the contiguous region in the projection image into spot candidates;

applying a filter to the spot candidates; and

counting the remaining spot candidates as spots.

59. The computer-readable medium of claim 58 wherein insignificant contiguous possible fluorescent image components are determined by comparing a size of a contiguous possible fluorescent image component with a threshold.

60. A computer-generated user interface for presenting results of microscopic observation of biological tissue subjected to a FISH experiment, the user interface comprising:

a scatter plot of sets of image components designated as spot candidates for the FISH experiment;

wherein the scatter plot comprises points indicating a size and intensity of spot candidates.

61. The computer-generated user interface of claim 60 wherein at least one point is operable to receive a user interface activation to navigate to a user interface display of information for a spot candidate associated with the point.

62. The computer-generated user interface of claim 60 wherein at least one point is operable to receive a user interface activation to navigate to a user interface display of a three-dimensional depiction of a spot candidate associated with the point.

63. The computer-generated user interface of claim 60 further comprising:

a display image depicting a view of the tissue subjected to the FISH experiment and a depiction of least one candidate spot thereon;

wherein the depiction of the candidate spot is operable to receive a user interface activation to designate the spot as a minimal intensity spot; and

wherein candidate spots designated as minimal intensity spots are visually emphasized when presenting the scatter plot.

SIGNAL COUNTING FOR IN SITU HYBRIDIZATION

Abstract

A computer system counts fluorescently tagged nucleic acid probe signals in biological specimens by determining a ratio of signals from a test probe to signals of a reference probe. Probe signals need not be counted with reference to cells, nuclei, or nuclear contours. Gene amplification or deletion can thus be detected by analyzing the ratio. Successive image slices are obtained by confocal microscopy, and the images are digitized. The digital images are transformed and analyzed to combine contiguous fluorescent signal segments in successive optical sections to identify discrete probe signals, or spots. Spots overlapping in the axial and transverse dimensions of a three-dimensional representation of the biological specimens can be distinguished. A graphical user interface presents various features for consideration by a user, who can provide guidance to a computer system counting the spots. Various features directed to identifying spot clusters and autofluorescent material can increase accuracy of spot counting.

10/088269

25/PRTS

JC10 Rec'd PCT/PTO 15 MAR 2002

SIGNAL COUNTING FOR IN SITU HYBRIDIZATION

FIELD OF THE INVENTION

This invention relates to methods of counting probe signals in biological specimens, such as probe signals produced by in situ hybridization in cells or tissue sections.

BACKGROUND OF THE INVENTION

Recent advances in molecular medicine have provided a greater opportunity to understand the genetic basis of disease, as well as the cellular mechanisms of disease, and select appropriate treatments with the greatest likelihood of success. Such diagnostic and prognostic cellular changes include the presence of tumor specific cell surface antigens (as in melanoma), and genetic abnormalities (such as activated oncogenes in tumors). A variety of techniques have evolved to detect the presence of these cellular abnormalities, including immunophenotyping with monoclonal antibodies, in situ hybridization with probes, and DNA amplification using the polymerase chain reaction (PCR).

One such technique for molecular diagnosis is in situ hybridization (ISH) in which labeled hybridizing agents (such as DNA, RNA, or single stranded or double stranded DNA probes) are exposed to intact tissue sections. The probes can be labeled by direct or indirect means. In direct labeling, the label (a chromophore) is integral to the probe. Direct labels include fluorescent dyes such as derivatives of rhodamine, fluorescein, and Texas Red, or enzymatic reporters such as horseradish peroxidase or alkaline phosphatase. Indirect labeling involves attaching a hapten (such as biotin, mercury or dioxygenin) to a probe through a linker. After hybridization, the hapten is detected using a labeled antibody or other specific binding protein.

When fluorescent dyes are used as labels, the technique is referred to as fluorescent in situ hybridization (FISH). Dyes such as fluorescein isothiocyanate (FITC) are often used to label a sequence of probe DNA in FISH. The probe hybridizes to a defined target nucleotide sequence of DNA in the cell, and the FITC fluoresces green when excited by a mercury arc lamp or Argon laser (in the case of

laser scanning microscopy), so that the labeled probe can be visually detected when the probed tissue is viewed through a microscope. Each chromosome containing the target DNA sequence will produce a fluorescent signal (sometimes called a spot) in every cell when the specimen is illuminated with suitable excitation. For example, specimens hybridized with a probe hybridizing to a specific region on chromosome 21 will produce two fluorescent signals in cells from normal individuals and three signals from Down's Syndrome (Trisomy 21) subjects who have an extra chromosome number 21. Alternatively, a fluorescent probe that hybridizes to an X chromosome will give one fluorescent signal per cell in males (who have only one X chromosome) and two fluorescent signals in females (who have two X chromosomes).

FISH is an excellent method for detection of gene copy number alterations in cancer and other diseases. FISH is a standard tool for analyzing congenital genetic alterations in clinical diagnostics (such as numerical chromosomal alterations, duplications, inversions and microdeletions). In cancer, characteristic gene amplifications or deletions are associated with the development and progression of the tumor. FISH analysis of gene amplifications in cancer can have prognostic and therapeutic significance, such as the detection of the HER-2 oncogene amplification in breast cancer. Detection of this oncogene by FISH was recently approved by the Food and Drug Administration as a diagnostic tool for human breast cancer.

However, a limitation on the more widespread use of FISH technology has been that counting of the fluorescent spots is extremely tedious, inaccurate, often highly subjective, and subject to substantial intra-observer variability. It also requires a highly trained technician who can recognize the cells or tissue to be analyzed and recognize and count tiny fluorescent spots accurately. Finally, at most 100 or 200 cells are typically analyzed per specimen, and in the case of gene amplification, much less than that (such as 20 per specimen). This can lead to statistical inaccuracy in defining the correct copy number.

A significant impediment to the accurate counting of fluorescent signals is that the probes hybridize throughout a three-dimensional nucleus, and the probe signals have to be counted from different focal planes for each nucleus. However, cells and probe signals can overlap in the two-dimensional view, and overlapping

signals are seen as a single signal. Such overlapping results in undercounting of signals, which can make it appear that an amplified gene is less amplified, or that fewer copies of a normal gene are present.

Given the tedium and subjectivity of signal counting, efforts have been made to automate this technique. For example, U.S. Patent No. 5,523,207 discloses a two dye FISH method in which probes are labeled with a first dye (such as FITC) and a contour of the nucleus is labeled with a second dye (such as propidium iodide or PI). These two dyes allow the number of signals per visualized nucleus to be determined. However, automated FISH spot counting using such techniques has been limited because FISH signals in the nuclei are often at different focal planes, resulting in interfering out-of-focus light. Moreover, automated detection of nuclear boundaries has been very difficult to perform in tissue sections. These factors have contributed to the inadequacy of existing algorithms for performing automated FISH.

It would therefore be helpful to provide a method and device for improving the accuracy of FISH spot counting.

SUMMARY OF THE DISCLOSURE

In one embodiment disclosed herein, fluorescently tagged nucleic acid probe signals are counted in a region of interest in a biological specimen by determining a ratio of signals from a test probe to signals of a reference probe, and the region of interest includes multiple cells. This is a contrast to prior approaches, in which probe signals have been counted with reference to cells or nuclei, and in which automated methods have counted probe signals with respect to stained nuclear contours.

In certain embodiments, the reference probe may be a fluorescently labeled polynucleotide (such as DNA or RNA) that hybridizes to the region of interest in a gene, and the reference probe is a polynucleotide labeled with a different fluorescent color, and which hybridizes to a reference target. The test probe may hybridize to a gene that is implicated or suspected to be involved in a particular disease, such as tumor development and progression. The reference probe may, for example, recognize a centromere of the same chromosome on which the gene of interest is contained. An increased ratio of test probe signals to reference probe signals would

then indicate an amplification in gene copy number, while a decrease in that ratio would indicate a relative loss of the gene of interest (such as a gene deletion). By determining a ratio between the quantities of test and reference signals, the problem of measuring changes in gene copy number with respect to a cell or nucleus is avoided. Typically, a count of centromeres approximates a nucleus or cell count.

Particularly accurate counting of FISH signals can also be accomplished by obtaining successive contiguous images of the region of interest to distinguish overlapping signals from the biological specimen. Without distinguishing the overlapping signals, they would otherwise obscure one another, and diminish accuracy of the spot count. In particular embodiments, the successive images are slices, such as digital microscopic optical sections from different depths of the biological specimen, obtained by confocal microscopy. The successive images are transformed by detecting and representing the positional values in each image of fluorescent emission signal segments, which make up the probe signal, as an array of digital values. Signal segments below a certain value (e.g. a threshold number of pixels) may be eliminated. The remaining digital signal segments may then be analyzed to combine contiguous fluorescent signal segments in successive optical sections into a single spot signal, which may be assigned to a particular optical section in which a strongest fluorescent signal segment is located, or a group of optical sections across which the contiguous signal segments have been detected. This localization of the fluorescent signal to a particular optical section allows overlapping spots to be distinguished, both in the axial and transverse dimensions of the three-dimensional representation.

The system detects the location of fluorescent spot signals in three-dimensional space by performing a morphological top-hat transform to digital images of the different levels to obtain fluorescent intensity spikes that indicate a spot signal segment. A threshold level of fluorescence intensity is determined to eliminate signal segments that are below a fluorescence intensity that would be associated with a valid spot signal. Remaining contiguous spot signal segments are segmented into a single spot signal, and the single spot signal may be assigned a location at a level associated with a greatest fluorescent intensity signal segment.

Certain disclosed embodiments also include devices for counting signals from in situ hybridization of probes in biological tissue, or determining a three-dimensional relationship between the signals, in which the device includes a confocal microscope, a digital camera positioned to obtain digital optical section images of different levels of the biological tissue, and a computer implemented system that detects and combines contiguous or adjacent signal segments (which are above a threshold) at the different levels, and separates vertically overlying, transversely overlapping, or non-contiguous signals from one another. After separating such signals from different levels into different spot signals, the computer implemented system then counts the spot signals, or compares their relative locations in three-dimensional space. Two or more different signal types (such as two or more distinguishable fluorescent dyes) may be used, one color for the test probe signal and a second color for the reference probe signal. The ratio of the number of test probe signals to reference probe signals can then be determined, or an unexpected overlap of the signals (as in a genetic translocation) can be assessed.

The test probe signals and the reference probe signals may be obtained separately, for example by successively illuminating the tissue specimen with light of different colors that selectively causes the different dyes to fluoresce, by viewing the specimen through filters that filter out signals (such as a filter that removes colors except the color of interest), or exposing separate contiguous tissue sections to the different probes. Alternatively, the test and reference probe signals can be obtained simultaneously, using multiple band-pass excitation and/or emission filters. Once the test and reference probe spot signals have been counted, a ratio of test probe signals to reference probe signals is calculated (without reference to boundaries of a cell or nucleus) to determine whether there is a genetic alteration, such as an alteration in gene copy number.

Alternatively, in certain disclosed embodiments, the FISH spots are accurately counted by obtaining the plurality of digital optical images at different levels of a biological sample, and constructing a three-dimensional image showing discrete fluorescent signal segments at different levels of the three-dimensional image. The three-dimensional image is constructed by determining a location of a fluorescent signal segment of a particular color in the different levels of the

biological sample, combining contiguous signal segments (which are above a pre-selected threshold) into a single spot signal, and separating non-contiguous signal segments into different spots signals. The location of signal segments in each level is determined by the presence of a fluorescent brightness intensity spike that indicates an increase in image component intensity as compared to a background intensity. The locations of signals of different colors can be similarly resolved in three-dimensional space, the number of spot signals of each color counted, and a ratio of the spot signals determined. This three-dimensional imaging would also allow one to study the orientation or location of the spots in the nuclei to make conclusions about the presence of genetic rearrangements that do not change copy number of the signals, such as translocations and inversions.

Certain features can be implemented to increase the accuracy of the spot counting. For example, under certain circumstances, a group of spots may be packed so closely together as to form a cluster that is difficult to analyze using a standard algorithm. Clusters can be automatically identified and counted using a cluster calibration feature, or a user can specify that a particular region of an image is a cluster.

Another feature for increasing accuracy filters out false spots by, for instance, eliminating spots appearing at a same location in the test and reference probe images or filtering out image data indicating autofluorescence. In addition, a user interface can be presented to allow a user to either select areas of particular interest for separate processing or specify areas not to be processed. In this way, the invention can benefit from information provided by a human operator.

In the copy number analysis, the ratio of signals can be a ratio of spot signals from a test probe that recognizes a gene of interest, and a reference probe that recognizes a chromosomal locus having an expected quantity in the biological specimen. The method can further include determining whether there is an increase in an expected ratio between the test signals and the reference signals, indicating an amplification of the gene of interest, or whether there is a decrease in the expected ratio between the test signals and the reference signals, indicating relative loss of the gene of interest.

This method is particularly applicable to high throughput techniques for performing automated FISH analysis of a large number of tissue, cell, or other specimens. The specimens can, for example, be in a tissue or cell microarray that includes specimens from the same or different sources,-such as tumors. In such embodiments, the method includes providing an array of biological samples, hybridizing the biological samples with a fluorescent test probe that hybridizes to a gene of interest in the biological samples and with a fluorescent reference probe that hybridizes to a chromosomal reference locus in the biological samples. Images are then obtained by confocal microscopy of contiguous sections at different depths of a plurality of the biological samples in the array, and fluorescent signal segments from the contiguous sections are detected. The contiguous signal segments in different sections are combined into a corresponding single spot signal, and the separate spot signals are resolved from one another. The distinct spot signals can then be counted.

In particular embodiments, the tissue array can include an array of many different tissue specimens, for example at least 50 tissue specimens, but hundreds or even thousands of tissue specimens can be included in the microarray. The tissue arrays can be constructed by obtaining a plurality of donor specimens, placing each donor specimen in an assigned location in a recipient array, and obtaining a plurality of copies of the recipient array in a manner that each copy contains a plurality of donor specimens that maintain their assigned locations. For example, a set of tissue samples relating to a same type of tissue from a plurality of donor specimens can be included. However, tissue arrays made by any method are suitable for use with the method of counting FISH spot signals.

The foregoing and other objects, features, and advantages of the invention will become more apparent from the following detailed description of disclosed embodiments which proceeds with respect to the accompanying drawings.

BRIEF DESCRIPTION OF THE FIGURES

FIGS 1A and 1B are schematic side and top views illustrating the problem of viewing three-dimensional FISH images in two dimensions.

FIG. 2A is a schematic view illustrating the prior approach of counting two color FISH spot signals with respect to cell nuclei (the contours of the nuclei being shown around the spot signals).

FIG. 2B is a view similar to FIG. 2A, but shows counting two color FISH spot signals and determining a ratio of the different colored signals in a region of interest (such as a field of view FOV) instead of with respect to cell nuclei. FIG. 2C is a photomicrograph of prostate tissue, illustrating a region of interest (ROI) within the tissue for purposes of calculating a ratio of test to reference probe signals.

FIG. 3 is a schematic view illustrating two color FISH, in which signal locations have been determined in three-dimensional space.

FIGS. 4A and 4B are flow diagrams illustrating one embodiment of an automated system for counting FISH spot signals in accordance with the present invention.

FIG. 5A is a flow diagram illustrating an algorithm for counting the FISH signals.

FIG. 5B is a series of optical sections 1-5, and a max image M, of prostate tissue subjected to two color FISH with a probe for the androgen receptor (labeled red) and a probe for the normal X chromosome (labeled green), in which red signals appear weaker than green signals.

FIG. 6A is a composite figure showing a print-out of optical sections of a series of eight confocal images (cuts 1-8), and a view of the image that would be seen (max-image) when the series of eight overlapping images would be viewed from above as a two-dimensional image.

FIG. 6B is a series of optical sections 1-5, and a max image M, of breast cancer tissue subjected to two color FISH with a ribosomal S6K probe (labeled red) and a probe for the chromosome 17 centromere (labeled green), in which red signals appear weaker than green signals. The excess of red S6 kinase signals over green reference signals indicates amplification of kinase.

FIG. 7 is a MATLAB graphical user interface (GUI) for displaying optical sections of a three-dimensional tissue specimen that has been subjected to FISH.

FIG. 8 is a histogram of the number of occurrences of the various gray levels, or fluorescence intensities, in the max image, which is used to threshold valid pixels and eliminate non-informative image components.

FIG. 9A is a print-out of a MATLAB GUI illustrating processing spot signal segments from different levels, and combining contiguous signal segments above a threshold into a single spot signal (which can be assigned to a particular level or levels) for purposes of counting a spot signal.

FIG. 9B is a schematic view of a vertical side view of optical sections, illustrating how image components below a certain threshold are eliminated to separate otherwise vertically contiguous signal segments, and separate overlying signals from one another.

FIG. 9C is a view similar to FIG. 9B, showing how transversely and vertically overlapping signal segments are separated by eliminating signal segments below a threshold, and grouping the remaining signal segments into discrete spot signals.

FIG. 10 is a three-dimensional representation of signal segments in different optical sections, resolved into a spot in an assigned optical section, and a combined image of the signals projected as spheres into a max image, in which a volume of the spheres is proportional to an intensity of the signal.

FIG. 11 is a view of a tissue section mounted on a slide, for processing by FISH.

FIG. 12 is a view of a tissue microarray mounted on a slide, for processing by FISH.

FIG. 13 is a flowchart showing an overview of an algorithm used to count FISH spots in a set of digital image slices.

FIG. 14 is a data flow diagram of an algorithm used to count FISH spots in a set of digital image slices.

FIG. 15A is an image showing FISH spots relating to the X centromere in normal prostate tissue, including false spots caused by autofluorescence.

FIG. 15B is a graph showing intensities of spot candidates for purposes of removing spots exceeding a threshold intensity.

FIG. 15C is a graph showing intensities combined with area for spot candidates for purposes of removing spots exceeding a threshold.

FIG. 15D is the image of 15A modified to emphasize spot candidates identified as autofluorescent tissue particles.

FIG. 16A is a scatter plot of spot candidates for an image relating to a FISH experiment conducted for centromere 17.

FIG. 16B is a scatter plot of spot candidates for an image relating to a FISH experiment conducted for the gene HER-2 and identifies spot clusters on the plot.

FIG. 17 is an illustration of a user interface for selecting areas of interest to be considered when calculating a spot count for an image.

FIG. 18 is an illustration of a user interface depicting a three-dimensional representation of a spot candidate.

DETAILED DESCRIPTION OF SEVERAL EXAMPLES

The present invention includes a method and apparatus for counting FISH spots, and particularly an automated, computer implemented method and device that helps avoid counting errors introduced by vertically or horizontally overlapping spots in two-dimensional projections of three-dimensional tissue sections. One of the problems solved by this invention is illustrated in FIG. 1A, which shows a vertical section through a stack of contiguous horizontal confocal layers of a tissue section that has been subjected to FISH. Hybridization probe signals are shown as spots in this figure, in which it can be seen in a vertical section that the spots are discrete and spaced along the z-axis. However, when viewed from above (in an x-y plane) as shown at the bottom of FIG. 1A, the discrete spots overlap and can not clearly be counted as separate spots. FIG. 1B illustrates a similar problem in which even more probe spots appear throughout the z-depth of the tissue section, but appear as an indistinguishable blur when viewed in an x-y plane from above, as shown at the bottom of FIG. 1B.

Another problem with the prior art is illustrated in FIG. 2A, which shows the conventional technique of counting hybridization signals within each cell nucleus. This view illustrates two color FISH, in which two probes are labeled with different dyes that fluoresce with different colors. The red label (R) may, for example, be

attached to a probe that hybridizes to a gene of interest (such as a hormone receptor gene that may be amplified in certain tumors). The green label (G), in contrast, may be attached to a probe that hybridizes to a known chromosomal locus that is not expected to vary in disease states (such as the centromere of a chromosome on which the gene of interest is found). The red label (R) is illustrated by a gray color in the schematic figure, while the green label (G) is represented by a darker color.

In a particular example, the gene of interest could be recognized by a probe for a gene on the X chromosome labeled with spectrum orange (to provide an orange-red spot signal), and the reference probe could be labeled with spectrum green (to provide a green spot signal) for the centromere of the X-chromosome. A single green signal (G) would therefore be observed in the nuclei of the schematic representation of FIG. 2A from male cells (which have only one X chromosome), while two green signals (G) would be seen in female cells. Amplification of the gene of interest would be noted in certain cells in the schematic representation of FIG. 2A in which there is an increase in the ratio of red signals (R) to green signals (G).

As shown in FIG. 2A, it is conventional to count the number of signals in each nucleus of a large number of cells. This approach has been adopted because amplification or deletion of a gene occurs in large populations of cells, and significant changes in copy numbers of genes are often only detected by examining a large number of cells (for example, at least 200). Because the amplification has been considered to be a nuclear event, a change in the copy number of a gene with respect to each nucleus has been counted, both manually and in automated systems. Such systems have been difficult to automate, however, because cells and nuclei overlap (as shown by the overlapping nuclear contours in FIG. 2A), and the nuclear contours have been difficult to reliably recognize in automated systems.

The present invention adopts a different approach which has been found to be more accurate, and has the additional advantage of being more accurately automated. This approach is shown in FIG. 2B, in which the ratios of probes are determined without reference to the cells (or the nucleus) in which the probes are contained. The nuclear contours highlighted in FIG. 2A are absent in FIG. 2B to illustrate this difference. Hence, FIG. 2B shows the FISH spots of FIG. 2A in a

region of interest (such as the microscope field of view [FOV] shown in FIG. 2A), but without reference to the nuclear contours illustrated in FIG. 2A. In accordance with the present invention, the ratio of test probes (R) to reference probes (G) in the region of interest is the ratio that is calculated. It has been found that this ratio provides sufficient information (over a sufficiently large number of cells in a region of interest) to be informative about the relative amplification or deletion of a gene of interest.

A region of interest is any arbitrary informative region across which an informative ratio can be determined. In some instances, a region of interest is a microscopic field of view at low magnifications (e.g. 100-200X). An entire microscope field of view can then be used for the image capturing and analysis at 400-1000X magnification (X40-100 objectives) for FISH analysis. The thickness of the tissue sections used for FISH analysis is the same as in sections routinely used for histopathological analyses, ranging from 4-10 μ in thickness.

Another example of a region of interest is an area which is selected for a specific analysis, as shown in FIG. 2C, which is a cross-sectional photomicrograph of tissue from a human prostate, as it would appear in a tissue microarray. In this figure, the region of interest (ROI) is the area of malignant cells that is outlined in black. A region of interest in this example is any homogenous site of a specimen, where most, if not all, specimens carry a particular alteration. However, contamination with non-altered normal tissue in the region of interest can be tolerated, if the copy number alteration in the abnormal cells is substantial.

In some embodiments, to further improve the ability of the system to accurately determine the ratio of test probe signals to reference probe signals, a view of the probe signals in three-dimensional space is constructed, as shown in FIG. 3. This view shows the three-dimensional relationship of the test and reference probe signals, which are interspersed among one another in all three dimensions of the tissue section which has undergone FISH. This three-dimensional view of the section can be further processed by a computer implemented system, as described below, to automatically count signals of each color, and to obtain a ratio of the different colored signals. This approach is particularly useful for high-throughput

analysis, for example, of tissue microarrays such as those disclosed in PCT publications WO US99/04000 and US99/04001.

In yet other embodiments, the three-dimensional imaging can be used to detect many kinds of genetic rearrangements in cells other than deletions or amplifications. For example, differently colored probes for the bcr and abl genes could be used to detect a fusion of these genes which occurs in a genetic translocation associated with chronic myelogenous leukemia (CML). See Tkachuk et al., Science 250:559-562, 1990. Such differentially labeled probes will flank translocation breakpoints, and produce a fusion after the translocation has occurred, but will be at separate loci if the translocation has not occurred. However, unless proximity or distance between the differentially labeled spots can be determined in three-dimensions, a measurement of the fusion will be inaccurate. Hence the method of determining a three-dimensional relationship of probe spot signals within a region of interest (such as a nucleus or other region) will permit a large scale quantitative analysis of three-dimensional distances between any two probe signals. For example, the three-dimensional coordinates of a red signal are determined, and the three-dimensional coordinates of a green signal are determined, and the coordinates are then analyzed to determine if they are overlapping (suggesting a translocation that has moved them into contiguous genetic loci) or separate (on different genetic loci than they would be normally).

EXAMPLE 1

Device for Performing Automated FISH

FIG. 4A shows such an automated spot counter 10 in accordance with the present invention. Briefly, the device 10 includes an automated optical microscope 12 (such as a confocal microscope) having a motorized stage 14 for the movement of a slide 16 relative to the viewing region of the viewing portion 18 of the microscope, a camera 20 for obtaining electronic or digital images from the optical microscope, a processing system 22 for counting the spots, and a memory 24 and a high resolution color monitor 26 for the storage and display respectively of images processed by the device 10.

In a disclosed embodiment, the classification device 10 is automated and computer implemented, and therefore also includes, in addition to the motorized stage 14, an automated apparatus for focussing, for changing lens objectives between high and low power, and for adjustment of the light incident of the slide, as well as circuitry for controlling the movement of the motorized stage, typically in response to a command from the processing system. The microscope may also include an automated slide transport system for moving the slides containing the specimen to be classified on to and off of the motorized stage, and a bar code reader for reading encoded information from the slide. An example of a microscope performing at least some of these functions is manufactured by Carl Zeiss, Inc. of Germany, or Atto Instruments of Rockville, Maryland. In particular embodiments, the microscope is a confocal microscope from Atto Instruments, such as that shown in PCT WO 99/22261, a Laser Scanning Microscope LSM 510 from Carl Zeiss, Inc., or an Axioplan 2 microscope from Carl Zeiss, Inc., equipped with a CARV module available from Atto instruments. An example of a camera 20 suitable for use with the invention, is a Quantix CCD camera available from Photometrics of Tuscon, Arizona.

The signal counting device 10 is shown in FIG. 4B with particular emphasis on the classification elements embodied in the processing system 22. The processing system 22 may include an image processor and digitizer 42, and a general processor 46 with peripherals for printing, storage, etc. The general processor 46 can be an INTEL PENTIUM microprocessor or similar microprocessor based microcomputer, although it may be another computer-type device suitable for efficient execution of the functions described herein. The general processor 46 controls the functioning and the flow of data between components of the device 10, may cause execution of additional primary feature signal counting algorithms, and handles the storage of image and classification information. The general processor 46 additionally controls peripheral devices such as a printer 48, a storage device 24, such as an optical or magnetic hard disk, a tape drive, etc., as well as other devices including a bar code reader 50, a slide marker 52, autofocus circuitry, a robotic slide handler, the stage 14, and a mouse 53. Although a single processor system is shown, the invention could also be carried out in a variety of other systems.

The depicted devices include computer-readable media such as a hard disk to provide storage of data, data structures, computer-executable instructions, and the like. Although the description of computer-readable media above refers to a hard disk, other types of media which are readable by a computer, such as removable magnetic disks, CDs, magnetic cassettes, flash memory cards, digital video disks, and the like, may also be used.

The image processor and digitizer 42 (hereinafter referred to as the image processor) digitizes images from the digital camera 20 and performs a primary algorithmic classification on the images to filter out unwanted information. The image processor 42 and the general computer 46 may each access read-only and/or random access memory, as would be readily apparent to one skilled in the art, for the storage and execution of software necessary to perform the functions described relative to that processing component. Further, each component 42 and 46 includes circuitry, integrated circuit chips, etc. for the control of communication or data transfer over the data bus 54, as well as other functions typical of similar processors.

The steps performed by the computer implemented system are illustrated in FIG. 5A. The tissue is first prepared and subjected to FISH (60) in the conventional manner, for example, as described in Example 2. The system then obtains multiple confocal optical sections (62), and the successive images are transformed by detecting and representing positional values in each image of fluorescent emission of a particular color as an array of digital values. Localized increases (spikes) in the digital values may represent signals in the optical section (64). A histogram is then generated (66) which produces a saddle point (described in association with FIG. 8) that permits fluorescent intensity image components below a threshold value to be eliminated (68, 70). Among the remaining image components, those that are substantially contiguous are grouped into segments according to various criteria (72). Additional filtering can be performed (74), for example, to account for autofluorescence and spot clusters.

With the spot signals then resolved in three-dimensional space, they are counted. Additionally, the counted signals can be combined into a single max image, which is similar to the two-dimensional view seen through a conventional

microscope. The max image is a convenient view for rapidly reviewing the FISH results in two dimensions.

In two-color FISH, signals from a second probe (usually of a different color) may also need to be detected (76). In that event, an optical filter can be changed to detect the new color, or an incident beam of laser light of a different color or wavelength can be directed at the tissue section. The signals of a different color are then resolved in three-dimensional space using steps 62-74 as previously described. Once no more signals of a different color are to be obtained, a ratio of the signals from the test probe to the reference probe is calculated to determine whether there is a gene copy alteration.

EXAMPLE 2

Example of Automated FISH Signal Counting From a Series of Confocal Images in High Throughput Analysis with Tissue Microarrays

Gene amplification is an important mechanism for the up-regulation of critical genes involved in cancer initiation and progression. A number of important oncogenes have already been found to be activated by DNA amplification. These include the HER-2 (17q12), C-MYC (8q24), PRAD1/CYCLIN D (11q13), FGFR-1 (8p12), and FGFR-2 (10q24) oncogenes. All of these are examples of genes for which alterations in gene copy number could serve as an indicator or disease onset or progression.

This example uses tissue microarrays, of the type shown in PCT publications WO US99/04000 and 04001 (which are fully incorporated by reference) as a high throughput technique for efficiently performing FISH on hundreds of tissue sample specimens. An example of such a tissue microarray is also shown in FIG. 12. In the absence of an automated technique for counting probe signals in the tissue microarray specimens, it would take many hours for each of the tissue specimens in the array to be examined and scored. However, using the automated technique described herein, hundreds or even thousands of the tissue specimens in the array can be examined and scored in much less time.

Prostate tumor microarray

Formalin-fixed and paraffin-embedded tumor and control specimens were obtained from the archives of the Institutes for Pathology, University of Basel (Switzerland) and the Tampere University Hospital (Finland). The least differentiated tumor area was selected to be sampled for the tissue microarray. The specimens that were interpretable for at least one FISH probe included the following: I) transurethral resections from 32 patients with benign prostatic hyperplasia (BPH) to be used as controls; II) 223 primary tumors, including 64 cancers incidentally detected in transurethral resections for BPH; 145 clinically localized cancers from radical prostatectomies, and 14 transurethral resections from patients with primary, locally extensive disease; III) 54 local recurrences after hormonal therapy failure, including 31 transurethral resections from living patients and 23 specimens obtained from autopsies; IV) Sixty-two metastases collected at the autopsies from 47 patients who had undergone androgen deprivation by orchiectomy, and had subsequently died of end-stage metastatic prostate cancer. Metastatic tissue was sampled from pelvic lymph nodes (8), lung (21), liver (16), pleura (5), adrenal gland (5), kidney (2), mediastinal lymph nodes (1), peritoneum (1), stomach (1), and ureter (1).

Construction and sectioning of tissue microarrays

The prostate tissue microarray was constructed using a tissue arraying instrument that created holes in a recipient paraffin block, and acquired tissue cores from the donor block by a thin-walled needle with an inner diameter of 0.6mm, held in a X-Y precision guide. The cylindrical sample was retrieved from the selected region in the donor and extruded directly into the recipient block with defined array coordinates. A solid steel wire closely fit in the tube was used to transfer the tissue cores into the recipient block. After the construction of the array block, multiple 5μm sections were cut with a microtome using an adhesive-coated tape sectioning system (Instrumedics, Hackensack, New Jersey). H&E stained sections were used for histologic verification of tumor tissue on the arrayed samples.

Fluorescence *in-situ* hybridization (FISH)

Two-color FISH to sections of the arrayed formalin-fixed samples was performed using a Spectrum Orange-labeled androgen receptor (AR) probe with corresponding FITC-labeled centromeric probes (Vysis, Downer's Grove, Illinois). The hybridization was performed according to the manufacturer's instructions. The following tissue treatment protocol was developed to allow formalin-fixed tumors on the array to be reliably analyzed by FISH: the slides of the prostate tissue microarray were first deparaffinized, acetylated in 0.2 N HCl, incubated in 1 M sodium thiocyanate solution at 80 degrees C for 30 minutes and immersed in a protease solution (0.5mg/ml in 0.9% NaCl) (Vysis, Downer's Grove, Illinois) for 10 minutes at 37 degrees C. The slides were then post-fixed in 10% buffered formalin for 10 minutes, air dried, denaturated for 5 minutes at 73 degrees C in 70% formamide/2x SSC (SSC is 0.3M sodium chloride and 0.03M sodium citrate) solution and dehydrated in 70, 80, and 100% ethanol, followed by proteinase K (4μg/ml phosphate buffered saline) (GIBCOBRL, Life Technologies Inc., Rockville, Maryland) treatment for 7 minutes at 37 degrees C. The slides were then dehydrated and hybridized. The hybridization mixture contained 3μl of each of the probes and Cot1-DNA (1mg/ml; GIBCOBRL, LifeTechnologies Inc., Rockville, Maryland) in a hybridization mixture. After overnight hybridization at 37 degrees C in a humid chamber, slides were washed and counterstained with 0.2μM DAPI. FISH signals were scored with a Zeiss fluorescence microscope equipped with a double-band pass filter using x40-x100 objectives. The relative number of gene signals in relation to the centromeric signals was evaluated for a region of interest. Criteria for gene amplification were as follows: at least 3 times more test probe signals than centromeric signals per cell in at least 10% of the tumor cells. Test/control signal ratios in the range between 1 and 3 were regarded as low level gains, and were not scored as evidence of specific gene amplification.

Output from the two color AR FISH is shown in FIG. 5B, which shows a series of optical sections 1-5, and a max image M that combines the fluorescent signal segments of the optical sections 1-5. A "signal segment" in an optical section, in combination with signal segments from other optical sections, are selectively chosen to make up a complete probe signal seen in the Max view. The red signal segments

in the optical sections (and the red signals in the Max view) have weaker signals than the green signals. Any of the two-dimensional optical sections 1-5, or the max image M, can be analyzed to determine the ratio of red to green signal segments or signals. As shown in FIG. 5B, the ratio of red to green (seen as a ratio of weaker to stronger spots) is substantially 1, throughout the optical sections 1-5 and in the max image M. This indicates that there is not an amplification of the AR, at least in the tissue section illustrated in FIG. 5B. In a gray scale version of FIG. 5B, the green signals appear as brighter spots, and the red signals are more difficult to detect.

Alternatively, a three-dimensional representation of the max image can be obtained by analyzing the optical sections that contain each colored signal segment, and separating spot signals that overlap in an axial (vertical) or transverse (horizontal) direction. The separated signals can then be counted, and a ratio of red to green signals more accurately determined. Examples of FISH output from the green labeled probe, obtained by the three-dimensional analysis system of the present invention, are shown in FIG. 6A, which illustrates FISH-signal counting from a series of confocal images. Cuts 1 through 8 are optical sections obtained by the confocal microscope filtered to obtain only the green signal segments at different levels, while the Max-image is a combination of the cuts 1-8 (and is similar to the superimposed images that would be obtained in conventional two-dimensional FISH). The total number of separate green signal segments counted in cuts 1-8 was 393 (summing the number of signal segments shown in brackets above each section number 1-8). After automated analysis of the signal segments, with separation of vertically and horizontally overlapping signals, it was determined that there were 283 unique signals. In a gray scale version of FIG. 6A, the green signals appear as slightly darker spots, such as the one jutting atop cut #6.

EXAMPLE 3

S6K Amplification in Breast Cancer

In this example, the biological consequences were examined of genomic rearrangements at 17q23, a locus amplified in up to 20% of primary breast cancers as assessed by comparative genomic hybridization (CGH). An array of primary

breast cancers was constructed, and used to determine S6K gene amplification frequencies in vivo.

Two cohorts of primary breast cancers were studied using the tissue microarray analyses. The first microarray consisted of 372 ethanol-fixed primary breast cancers. The second microarray consisted of 612 primary breast cancers from the years 1985-1995, from patients with complete clinico-pathological information, including an average of 5.4 years of follow-up. Both series of tumors were analyzed, with 668 cases being informative for all experimental and clinical data. Both tumor cohorts were obtained from the Institute of Pathology, University of Basel. The tumor samples included 73.3% ductal, 13.6% lobular, 3% medullary, 2.6% mucinous, 1.5% cribriform, 1.4% tubular, 1.1% papillary carcinomas, 1.9% ductal carcinoma in situ, and 1.7% of other rare histological subtypes. The grade distribution was 24% grade 1, 40% grade 2, and 36% grade 3. The pT stage was pT1 in 32%, pT2 in 51%, pT3 in 7%, and pT4 in 10%.

A SpectrumOrange labeled PAC probe specific for S6K and a SpectrumGreen labeled chromosome 17 centromere probe (Vysis, Downers Grove, IL) was used for copy number analysis. Interphase FISH to breast cancer cell lines was done as previously described in Barlund et al., *Genes Chrom. Cancer* 20:372-376, 1997. The hybridizations are evaluated using a Zeiss confocal fluorescence microscope, and following the algorithm shown in FIG. 5A.

The breast cancer tissue section showed a high-level SK6 gene amplification, with a higher number of red signals than green reference signals, as shown in FIG. 6B. In a gray scale version of 6B, the red signals appear as brighter dots, such as the loose cluster of dots in the center of image 5.

EXAMPLE 4

Graphical User Interface for Three-Dimensional FISH Signal Counting

Although the invention can be implemented in a variety of computing environments, the following examples are implemented in MATLAB, which is available from Mathworks of Natick, Massachusetts. FIG. 7 shows a graphical user interface (GUI) that is displayed by the system for processing the fluorescent images and counting signals associated with probes that have hybridized to a target nucleic

acid sequence. The image- displayed on the interface is an image of one of the optical sections of a tissue section that has undergone FISH (in this case slice=5 refers to the 5th of 8 contiguous optical sections at successively deeper depths of the tissue section; stack=7 refers to the number of the tissue section on the tissue array that is being processed; and threshold=26 refers to a threshold value of signal intensity below which intensity values are eliminated from calculations).

FIG. 8 is a histogram which illustrates how the threshold value (threshold=26) is determined in this example. The histogram plots the relative frequency of gray levels (which corresponds to signal intensity) across the pixels of the image of each optical section, where the x-axis is the frequency of occurrence, and the y-axis is a brightness value. This graph shows two modes, in which the first mode (the sharp spike) represents noise (primarily dark background), and the second mode (the broader peak) represents useful signals (such as brighter pixels associated with fluorescent intensity). To determine a saddle point (threshold value), the histogram is smoothed, and the derivative of the resulting graph is calculated. The point T=26 is the point at which the derivative of the curve changes value from + to -, and this is the value that is selected as the threshold. Pixels having a brightness below this level are eliminated from further data manipulation.

FIG. 9A shows an image which is displayed of contiguous signal segments in different optical sections 1-8 of a tissue stack (which corresponds to a microarray spot). This FISH signal is designated spot 224, and the image intensity of the signal segments that are present on optical sections 1-8 can be seen in the small image panels across the bottom of the display. Each small panel 1-8 is a section of the three-dimensional space within the tissue section, and the number of pixels where signal is present is shown by the white boxes in the display panels. Although several three-dimensional representations of the signal segments are shown in FIG. 9, the center representation (with segments labeled 1a-8a) will be used for purposes of illustration to explain how the different signal segments are combined into a single spot signal, or how certain signal segments are discarded.

The relatively small brightness signals (which correlate with respectively few white pixels) in panels 1, 2, 3 and 4 are mapped into correspondingly small three-dimensional geometric boxes 1a, 2a, 3a and 4a, having a volume proportional to the

area of the illuminated pixels in the panels 1-4. The absence of any illuminated pixels in panel 5 corresponds to an absence of any geometric figure in space 5a (or can indicate that any pixels were below a threshold value and therefore not considered for further data manipulation). The area of the white pixels in panel 6 correspond to the volume 6a, and the area of the white pixels in panel 8 corresponds to the volume 8a, while the very large illuminated area in panel 7 corresponds to the figure 7a, which has the largest volume of any of the geometric figures.

One function of the algorithm is to eliminate background noise from actual probe signals. This can be done by setting a threshold value, below which a signal segment can be considered not to be a valid signal segment. The signal segments in panels 1, 2, 3, 4 and 6 do not meet the threshold criteria for minimum illuminated pixels, and are therefore not considered when combining the signal segments into an actual signal. The signal segments in panels 7 and 8 do, however, meet the criteria for minimum pixel illumination. Since more pixels are illuminated in panel 7 than in panel 8, the location of the signal is assigned to level 7, although alternatively it could be considered to reside in both levels 7 and 8.

Other signal overlap situations that the algorithm can resolve are also shown schematically in FIGS. 9B and 9C. In FIG. 9B a side view of optical sections 1-10 is shown, in which two spot signals vertically (axially with respect to the microscope lens) overlap one another. The signal segments in sections 1, 5, 6 and 10 fall below a pre-selected threshold value, and are eliminated from further consideration. The remaining signal segments in sections 2, 3, 7, 8 and 9 are above the threshold, and are grouped into a first spot signal at contiguous sections 2-4, and a second spot signal at contiguous sections 7-9. The probable location of the top spot signal can be assigned to section 3, where the strongest signal segment of the top spot signal is found, or alternatively the spot signal can be considered to reside at sections 2-4. The probable location of the bottom spot signal can be assigned to section 8, where the strongest signal segment of the bottom spot signal is found, or in sections 7-9 across which the signal segments extend.

FIG. 9C schematically illustrates how vertically and/or horizontally (transversely with respect to the optical axis of the microscope lens) overlapping spot signals can be resolved. A side view of optical sections 1-8 is shown in which

the signal segments in sections 1 and 8 do not meet a threshold, and are eliminated. In section 4, there are two horizontally non-contiguous signal segments. The signal segment 4B (represented by a gray bar) does not meet a threshold value and is eliminated, while the segment 4A (represented by the black bar) does satisfy the threshold and remains for consideration. Similarly, there are two non-contiguous signal segments in section 5. The segment 5B (represented by the gray bar) meets the threshold, while the segment 5A (represented by the black bar) does not meet the threshold and is eliminated. Once the signal segments 4B and 5A are eliminated, the segments can be resolved into a top spot signal (comprising segments 2, 3 and 4A) and a bottom spot signal (comprising segments 5B, 6 and 7)

FIG. 10 illustrates a three-dimensional representation of a tissue section that has been subjected to FISH and analyzed by the method of the present invention. The three-dimensional space is divided into x-y-z coordinates, in which the z axis is associated with successive optical sections 1-8 of the tissue section, and the signal segments of all of the signals in the section are illustrated. Signal segments are illustrated as small black cylinders, and the assigned location of each spot is illustrated as a gray colored larger cylinder, some of which are designated A1, A2, A3, A4, A5, A6, A7 and A8. Each of these gray cylinders A is associated with a corresponding sphere B in the max plane, in which the volume of each sphere B can be proportional to the volume of the A cylinder with which it is associated.

The system additionally provides an opportunity for a user to provide guidance during spot counting. For example, the user can specify a particular area of interest by selecting it on the screen. Typically, the max image (e.g., image M of FIG. 5B) is presented, and the user may select an area or areas via a pointing device (e.g., a mouse). Counting is then limited to only the selected area or areas. Such a feature can be particularly useful when the user recognizes that a certain area of the image relates to a region of interest.

The system also provides a way to eliminate a specified area or areas selected via a pointing device (e.g., a mouse). Portions of the image within the specified area or areas (sometimes called "gated areas") is ignored when spots are counted.

EXAMPLE 5

Algorithm Overview

The method of the present invention facilitates the analysis and quantitation of FISH on tissue microarrays, by overcoming previous obstacles to automation. This method, as implemented in the algorithm of this example, applies dual-color FISH with different gene probes and corresponding chromosomal reference probes to tissue microarrays. Because each tissue section is a three-dimensional volume, a confocal microscope system is used to generate a three-dimensional stack of a number of fluorescent images (e.g., 24 images) over defined tumor areas of each of the 4 μm thick tissue specimens on the tissue microarrays. The algorithm identifies signal segments on each level along the Z-axis, and differentiates signal segments which overlap vertically above the X-Y plane.

The algorithm can be refined to account for spot clusters, and a variety of filtering mechanisms can be used to remove signals related to, for example, auto-fluorescence and other false signals. A variety of parameters used during the algorithm can be modified by the user to account for circumstances related to a particular image or set of images. For example, since the magnification used for an image can vary, the parameters of the algorithm can be modified to account for the relative size of a pixel. There are many other examples of parameters (e.g., numeric values) that can be varied to avoid a strict adherence to a particular value. Various results of intermediary calculations are stored to facilitate efficient re-calculation when parameters are modified by the user.

An overview of an implementation of the algorithm is shown in FIG. 13. Typically, the algorithm is provided with a set of digital image slices (e.g., a raster image) representing a set of observations (e.g., of a biological specimen subjected to FISH) taken at different depths along a z-axis via a confocal microscope. Thus, the set of images is sometimes called a "stack." The stack typically consists of anywhere from 8-32 images, although more images may be used. Satisfactory results have been obtained with 24 images.

At 1302, possible fluorescent image components are identified in the slices. For example, pixels meeting certain criteria are selected as possibly corresponding to a location within or near a spot and placed into a set of binarized images. The

criteria can include an intensity level, and the images may be subjected to various filters before constructing the binary image.

At 1304, the resulting pixels in each slice are projected onto a projection image. For example, a binary projection image can be constructed from a stack of binary images by setting pixels in the binary projection image at locations (X,Y) whenever there is any pixel set in any of the stack of binary images at the same (X,Y) location (regardless of the Z location).

At 1306, insignificant regions in the image slices are discarded. So, for example, single lone pixels in a slice, or a contiguous set of 2-3 pixels are ignored, removed, or labeled as insignificant. A similar process can be applied to the projection image.

At 1308, each contiguous region in the projection image is considered in turn. For example, minimal rectangles can be drawn around each substantially contiguous region in the projection image, numbered, and considered in turn.

At 1310, contiguous regions in the image slices associated with the projection image region under consideration are analyzed and grouped into spot candidates. At 1320, the next region in the projection image is considered, until each has been considered.

At 1322, filtering an augmentation can be applied to the candidate spots. For example, certain candidate spots may be false spots, or others may actually be clusters representing plural spots.

Finally, at 1324, the spots are counted. The spot count can then be used vis-à-vis other spot counts to provide a ratio (e.g., a ratio of genes to centromeres).

EXAMPLE 6

Details of Algorithm

In this method, the traditional signals-per-nucleus approach is replaced by an overall gene-to-chromosome ratio in a defined tumor area. Therefore, the ratio between the signals of gene probes and the reference probes are calculated irrespective of the number of nuclei. This new method for automated capturing and quantitation of FISH signals significantly advances the utility of tissue microarrays

as a high-throughput tool for detection of relative gene copy number alterations during cancer development and progression.

Although the invention can be implemented in a variety of computing environments, the following examples are implemented in MATLAB, which is available from Mathworks of Natick, Massachusetts. Various MATLAB-based Graphical User Interface (GUI) tools allows visualization of three-dimensional shapes of spots as well as their two-dimensional projections on the slices. The dynamic interface provides for manipulation with many input parameters for the threshold, number of slices, and filtering. It also provides for visualization of differently colored spots, storage, displaying and printing data in the form of different 3-D images, diagrams, and tables.

An implementation of the algorithm in a scenario with 24 image slices can have the following features:

(1) A morphological top-hat transform is applied to each of the 24 images to yield 24 outputs, each possessing brightness intensity spikes jutting above an essentially flat background. Each bright spike can correspond either to a signal segment, or to noise.

(2) Each top-hat image is thresholded to produce a stack of 24 binary images showing spike locations. Morphological filters are applied to the binary images to eliminate noise, and touching spots are segmented.

(3) Binary spot markers occurring as vertical neighbors in the stack are grouped into one final spot located at a particular assigned stack level. Various parameters of the algorithm are set to fit the physical characteristics of the images. These include window size for the top-hat transform, threshold levels, and sizes of filter structuring elements.

The performance of the algorithm can be based on the following main steps:

1. Given $L + 1$ slices of the fluorescent images, $\boldsymbol{X_0}$, $\boldsymbol{X_1}$, ..., $\boldsymbol{X_L}$, of sizes $N\times M$, calculate the max-image $\boldsymbol{Y_{max}}$:

$$Y_{\max}(j,i) = \max\{X_k(j,i); k = 0,1,\ldots, L\}.$$

The histogram, $H(t)$, of this image is calculated and smoothed, and then the approximate value of the threshold, T, is determined as the minimal saddle-point between two modes of the histogram. The histogram is displayed on the screen and the threshold can be changed on-line. The chosen value of the threshold is used in the next steps for computing binary images.

2. For each image X_k, $k=0,1,\ldots,L$, the top-hat transform is calculated. Sometimes the top-hat transform is called "image minus opening." The calculation is

$$Y_k = X_k - X_k \circ B, \qquad (1)$$

where $X_k \circ B$ is opening by the structuring element B, which is taken to be a circle fitting within a 7x7 square (roughly, a circle with a diameter of 7). Other structuring elements can be used (e.g., the following examples show a 5×5 square) and can be specified by the user via a user interface feature. The opening is defined by $X \circ B = (X \varnothing B) \oplus B$, and calculated as follows:

$$(X \varnothing B)(j,i) = \min\{X(j + j_1, i + i_1); (j,i) \in B\}, \qquad (2)$$

$$(X \oplus B)(j,i) = \max\{X(j + j_1, i + i_1); (j,i) \in B\}. \qquad (3)$$

To calculate the top-hat transform in (1), the decomposition of the transform by two top-hat transforms with 1×5 and 5×1 structuring elements is used, which results in the fast performance of the transform 5×5:

$$X \varnothing \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} = \left(X \varnothing \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right) \varnothing \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix}, \qquad (4)$$

$$Y \oplus \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} = \left(Y \oplus \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right) \oplus \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \qquad (5)$$

The data of the top-hat transform $\boldsymbol{Y_k}$ is stored for future processing in case the value of the threshold needs to be changed.

3. Each top-hat transform $\boldsymbol{Y_k}$ is thresholded by the value T, resulting in a binary image $\boldsymbol{B_k}$, $k=0,1,\ldots,L$,

$$B_k(j,i) = 1, \text{ if } Y_k(j,i) > T,$$
$$B_k(j,i) = 0, \text{ if } Y_k(j,i) \le T.$$

A projection of multiple slices of the image can be calculated as

$$B_{\max}(j,i) = 1, \text{ if } B_k(j,i) = 1,$$

for at least one k. Typically, the projection is calculated with reference to all available image slices, in which case the binary image $\boldsymbol{B}_{\max}$ is sometimes called the "projection binary image" because it is roughly equivalent to peering down through the stack of binary images. In some sense, the binary image $\boldsymbol{B}_{\max}$ is also a "maximum" image; however, in the example, the maximum is computed comparing only 1's and 0's.

4. The obtained binary images $\boldsymbol{B_k}$, are processed to find the separate regions revealed by the thresholded top-hat transform. Each actual spot is composed of some number of these regions on different binary slices. Each of these slice regions is referred to as a signal segment, so that each spot signal is composed of a union of signal segments. The software can identify signal segments by locating substantially

contiguous groups of pixels, number the signal segments, and calculate the following data:

A. coordinates of the signal segment center;

B. the number of pixels composing the signal segment;

C. the coordinates of the minimal rectangle which contains the signal segment.

Signal segments composed of a single pixel are ignored during this procedure but their presence can be used during other steps.

5. The projection binary image $\boldsymbol{B}_{\max}$ is processed similarly, and its separate regions (projection binary image regions) are identified by locating substantially contiguous groups of pixels and numbered. The following are calculated for spot signals on the projection binary image:

A. coordinates of the projection binary image region center;

B. the number of pixels (1's) composing the region;

C. the coordinates of the minimal rectangle which contains the region;

D. the number of signal segments from the different optical sections, which lie inside the surrounding rectangle of the region; and the largest such signal segment is marked.

The algorithm then attempts to identify each of the spots by integrating the signal segments into spot signals via analysis of the sizes, intensities, and locations (in 3 dimensions) of the signal segments. Typically, for purposes of determining whether a particular signal segment is of sufficient size to indicate a contiguous connection, a threshold (e.g., 3 pixels) is useful to designate as a cutoff. So, for example, signal segments having 3 or fewer pixels can be ignored when determining whether the slice above and/or below a signal segment is contiguous with the particular signal segment.

Specifying such a threshold can be particularly useful in resolving closely-neighboring spots, such as two spots having substantially similar (X,Y) coordinates but appearing in different slices of the image. Examples of such scenarios are shown in FIGS. 9B and 9C. As shown, there can be more than one spot in a vertical direction or in a horizontal direction.

During processing, segments that are substantially vertically contiguous can be associated with a particular optical slice having the greatest signal intensity, which is often the center of a single spot. However, it may be that there is another image slice with great signal intensity in the same segment, in which case there may be more than one spot in the segment.

Typically, the information about a spot candidate includes a size (e.g., total number of pixels across the image slices) and an intensity (e.g., the brightest pixel in the region of the slice having the most pixels for the spot candidate).

A data flow diagram for an implementation of the algorithm is shown in FIG. 14. The slice data 1402 is used to construct a max image 1404 and the top-hat slice data 1406. The max image 1404 is useful for presenting to the user to give the user an informative depiction of a combination of the image slices and for various user interface features. It can be constructed by taking the pixel with the greatest intensity out of the image slices in the z-plane.

The top-hat slice data 1406 can be generated as described above. Max top-hat image 1408 can be generated by finding the greatest intensity of the top-hat slices for each (X,Y) location. The max top-hat image 1408 need not be displayed, but is useful, in combination with the top-hat slice data 1406 to construct the histogram 1410.

A threshold 1412 is chosen with reference to the histogram 1410 as described above. Then, the thresholded slice data 1420 can be constructed by generating a binary image for each of the slices, where pixels in the slices are set if the slice image in the slice data 1402 exceeds the threshold 1412.

From the thresholded slice data 1420, substantially-contiguous regions of the thresholded slice data can be identified to generate regions 1422. The regions are filtered to produce filtered regions 1424 of the thresholded data.

Also from the thresholded slice data 1420, a projected binary image 1426 can be constructed, and substantially-contiguous regions 1430 within the projected binary image 1426 found. These regions can be filtered to generate filtered regions 1440 of the projected image.

Z-axis analysis data 1428 can be generated with reference to the filtered regions 1424 and the regions 1430 and used to generate the candidate spots 1432,

from which, false spots 1434 and clusters 1442 can be identified. Additional filtering can be performed to generate spot count 1444. Finally, display data 1450 can be used to present graphical results to the user.

The data flows in the diagram are general only, and often data in upstream data components can be relied upon by downstream data components (or vice versa), even if not specifically shown in FIG. 14. As the data components are generated, they are often saved to permit efficient re-calculation when a parameter is adjusted to account for a particular image scenario.

In addition to the above-described filters, additional filtering mechanisms can be provided to improve accuracy of spot counting. For example, one feature relates to removing false spots relating to, for example, autofluorescence. Two separate approaches relate to removing based on intensity and intensity in combination with area. For example, after candidate spots are identified, an intensity of the spot can be calculated. If the intensity exceeds a certain threshold (say, two standard deviations above the mean intensity), the spot can be discarded and not counted. In addition, or alternatively, the intensity can be combined with (e.g., multiplied by) the area or volume of a spot. Again, if the value is over a particular threshold, the spot can be discarded and not counted. The approaches can be combined.

For example, the image in FIG. 15A illustrates spot candidates in an image of normal prostate tissue in which the X centromere has been labeled via FISH. Some of the spot candidates are actually autofluorescent tissue, and can be eliminated from the spot count. The image can be presented as part of the user interface.

The graph in FIG. 15B shows the intensity of the 149 spot candidates identified in the image. A mean and standard deviation of the spot candidate's intensities is calculated, and used as a threshold (e.g., two standard deviations above the mean). Spot candidates having intensities above the threshold are discarded (e.g., identified as "small autofluorescent tissue particles").

The graph in FIG. 15C is similar to that of 15B, but also includes an area component (e.g., by multiplying the intensity by the number of pixels" and not included in the spot count). Again, spot candidates having a rating above the

threshold are discarded (e.g., identified as "large autofluorescent tissue particles" and not included in the spot count). Eight candidates can be removed because they exceed the threshold. The threshold in either case can be adjusted manually by a user if the default setting is inappropriate.

FIG. 15D shows the image of 15A, but the small autofluorescent tissue particles, large autofluorescent tissue particles, and the true FISH signals are differentiated by presenting each as a different color. The image can be presented as part of a user interface for the system. Eliminating the autofluorescent particles typically leads to more accurate spot count results. In a color version of the image, the image components 1510, among others, are portrayed in the color green to indicate that they are considered true FISH signals. The image components 1520 and 1521, among others, are portrayed in the color yellow to indicate that they are small autofluorescent tissue particles. The image components 1530-1533, among others, are portrayed in the color blue to indicate that they are large autofluorescent tissue particles. A variety of other colors or other ways of emphasizing and differentiating the image components related to autofluorescence can be used.

Still another feature than can be implemented is removal of spots appearing at substantially identical positions in two channels (e.g., red and green channels generated in two separate images, such as an image for a test probe and an image for a reference probe). Typically, such a situation means the spot can be ignored (e.g., not counted for either image).

Yet another feature relates to identifying clusters and estimating the number of spots in a cluster. In some cases, certain criteria (e.g., the size of a projection binary image region or the total number of pixels in slice regions related to a projection binary image region) indicates that a cluster of spots is present. And, the software can be configured to accept user guidance as to which portions of the image are a cluster (e.g., via selection by pointing device). Determining the number of spots in a cluster can be difficult; however, one way of providing accurate estimates is to provide a calibration mechanism.

For example, in one instance, 200 sets of clusters were analyzed manually to determine how many spots were in the clusters. These clusters were then subjected to the above-described algorithmic analysis. Under certain circumstances, a factor

of 2.5 was determined to be appropriate for determining how many spots were in a cluster. Thus, for example, if an area is identified as a cluster and the standard algorithmic analysis shows 4 spots, an estimate of 10 is provided. Other, more complex, analysis can be done, such as, for example, counting the total number of pixels associated with the signal segments related to a projection binary image region.

For clusters, a mapping between spots detected and actual spots can be used. Sometimes detection of a small number of spots (e.g., 3) may in fact indicate a tight cluster of (e.g., 9) spots, while detection of a larger number of spots (e.g. 7) may be a loose cluster of (e.g., 9) spots. Therefore, a uniform gain factor is not appropriate. The appropriate method for handling clusters may vary depending on the target being counted. For example, certain genes have a higher propensity for clustering than do centromeres. The user can manually adjust the cluster detection and counting parameters manually.

The number of spot signals is counted and the results are provided to the user. The interactive options of the program allow the threshold *T* to be changed, and process the spot counting of the fluorescent images from step **3**, avoiding the repeated calculation of the top-hat transforms.

The output data using this algorithm can include:

1. Original images of all slices.
2. Max-image.
3. Histogram of Max-image.
4. Max-image with colored spots.
5. Top-Hat transforms of the images.
6. Number of spots.
7. Max-image with colored spots after filtering out spots with less than some specified number of pixels.
8. Original images with colored spot-cuts.
9. Images with colored rectangles surrounding the spots.
10. 3-D view of each spot with sections lying on slices by vertical.
11. 2-D view of spot composition by projections.

12. 3-D view of a number of spots with their projections by slices.

13. Table of data for spot-cuts for L slices, which consists of:

# spot-cut	# slice	x_{center}	y_{center}	# pixel	x_{left}	x_{right}	y_{bottom}	y_{top}	r.v.l.
24	4	160	124	24	160	164	120	125	1

where r.v.l. is the abbreviation of relative vertical location, x_{left}, x_{right}, y_{bottom}, and y_{top} are the coordinates of the rectangles surrounding the spot-cuts. The last box indicates whether the spot-cut is below, on, or above the true spot.

14. Table of data for spots on the Max-image, which consists of:

# spot-cut	x_{center}	y_{center}	# pixel	x_{left}	x_{right}	y_{bottom}	y_{top}
224	182	164	42	180	186	160	167

EXAMPLE 7

Scatter Plot of Spot Candidates and Calibration Features

An additional user interface can be presented for assisting in analyzing and manipulating the data related to FISH experiments. For example, FIGS. 16A and 16B show scatter plots generated for a FISH experiment in which the presence of centromere 17 and gene HER-2 were detected to generate a ratio indicating possible amplification of HER-2.

Spots relating to centromere 17 typically do not cluster and are shown in FIG. 16A. Spot candidates are represented by appearing in a position appropriate for their intensity and size (e.g., number of pixels). The user can set a threshold parameter (e.g., 4) so that candidate spots having fewer than the specified number of pixels are excluded from the spot count. Also, a minimum intensity parameter can be set.

Spots relating to HER-2 are shown in FIG. 16B and are more difficult to count due to potential clustering. For purposes of illustration, areas of the plot roughly relating to background noise 1652, real signals 1654, and clusters 1656 are circled. To facilitate determining an appropriate threshold, the user is presented with a calibration feature. An image (e.g., the maximum image 1404) is presented to the user, who can then verify that certain spot candidates are indeed spots,

although there are of minimal intensity. The user selects the spot candidate shown and activates a user interface element indicating the spot candidate is of minimal intensity.

Subsequently, when the scatter plot 1650 of FIG. 16B is shown, the spot candidates of minimal intensity are shown in red to facilitate choosing an appropriate minimum intensity threshold (e.g., by adjusting an appropriate parameter in the algorithm), below which spot candidates are not counted as spots. Thus some of the points 1658, typically near the bottom of the plot, are red. The minimum intensity threshold is typically adjusted to barely include the minimum intensity candidates.

Also useful is the illustration of clusters 1656, which can be used to adjust the cluster selection criteria. Additionally, the user can at any time click on a particular spot candidate (e.g., 1660) to navigate to information about the spot candidate, including a three-dimensional view of the spot candidate. At the user interface, the user can then manually designate the spot candidate as a cluster or non-cluster.

The scatter plot 1650 thus presents a useful tool for allowing the user to more easily grasp the totality of the image data and easily navigate to appropriate user interface screens for analyzing and adjusting the data, as well as calibrating the system.

EXAMPLE 8

User Direction of Areas to be Analyzed

The system can also provide a user interface by which the user can guide the selection of areas to be analyzed. Although the system typically discards clearly aberrant data (e.g., a substantially total black or high-intensity image slice), the system may benefit from assistance from the user in determining which areas of an image are of interest or which are not. As described below, the user can select an area to be counted separately or an area not to be counted.

FIG. 17 shown a user interface 1702 presented to the user for designating areas of interest. Typically, the user interface 1702 presents a max image (e.g., 1404). By using a pointing device (e.g., a mouse or trackball), the user can designate a particular area 1706 as one to be processed separately. Multiple areas

can be added to an area set and processed separately. In this way, the user can direct the system to focus on particularly relevant portions of the image.

The system also supports designating areas as not to be processed or included in the spot count. The algorithm can store the areas in a list and, for example, not count spots at a location (e.g., X, Y coordinates) within the area(s) designated.

The above feature is useful because areas of the image can sometimes be identified as debris, stroma, connective tissue, or blood vessels, which tend to interfere with determining an appropriate spot count in certain scenarios (e.g., some material may be particularly prone to autofluorescence).

EXAMPLE 9

User Interface Including Three-dimensional Representations of Spot Candidate

Still another user interface feature can be presented by the system to assist in evaluation and manipulation of the image data. FIG. 18 shows a user interface 1800, which depicts a spot candidate in three views, 1802, 1804, and 1806. The views assist in determining, for example, whether the spot candidate is a cluster, as is likely the case in the illustrated example. The view 1806 is additionally processed to give a smoother representation of the spot candidate. The horizontally-contiguous region with the greatest area (e.g., number of pixels) is specially identified in the views with emphasis 1808. A status line 1818 indicates information about the spot candidate.

A strip along the interface indicates the max image, and the 15 image slices (numbered 0-14) in binary form after thresholding. The image slice having the greatest area region is noted at 1820. User interface controls 1832 and 1834 allow navigation to spot candidates and identified clusters, respectively.

In view of the many possible embodiments to which the principles of the invention may be applied, it should be recognized that the illustrated embodiments are examples of the invention, and should not be taken as a limitation on the scope of the invention. Rather, the scope of the invention is defined by the following claims. We therefore claim as our invention all that comes within the scope and spirit of these claims.

We claim:

1. A method for counting nucleic acid probe signals in a region of interest in a biological specimen, the method comprising:

counting a number of test signals from a test probe;

counting a number of reference signals from a reference probe; and

determining a ratio of the counted test signals from the test probe to the counted reference signals from the reference probe, wherein the region of interest comprises multiple cells.

2. The method of claim 1, wherein the reference probe is a polynucleotide that hybridizes to a centromere, and the number of reference signals from the reference probe approximates a nucleus count in the biological specimen.

3. The method of claim 1, wherein the reference probe recognizes a target on a same chromosome as the test probe.

4. The method of claim 1, wherein the test probe is a polynucleotide that hybridizes to a target sequence in a gene, and the reference probe is a polynucleotide that hybridizes to a reference sequence.

5. The method of claim 3, wherein the reference probe recognizes a centromere of the same chromosome on which the gene of interest is contained.

6. The method of claim 1, further comprising obtaining successive images of the region of interest to distinguish overlapping signals in the biological specimen.

7. The method of claim 6, wherein the successive images are optical sections of the region of interest.

8. The method of claim 7, wherein the optical sections are at different depths of the biological specimen.

9. The method of claim 8, wherein the successive images are transformed into digital representations in which contiguous signal segments in successive optical sections are combined into a single signal in a particular optical section in which a strongest signal segment is located.

10. The method of claim 6, wherein different successive images are obtained for the test probe signals and the reference probe signals, and a quantity of test probe signals and.reference probe signals are determined.

11. The method of claim 6, wherein successive images are obtained which show distinguishable test probe signals and reference probe signals, and a quantity of the test probe signals and reference probe signals are determined.

12. The method of claim 6, wherein the successive images are obtained by confocal microscopy.

13. The method of claim 1, wherein the ratio of signals is determined without reference to boundaries of a cell nucleus.

14. The method of claim 1, wherein the ratio of signals is determined without reference to the boundaries of a cell.

15. A method of counting visible signals from probes used with in situ hybridization of biological samples, the method comprising:

obtaining a plurality of images at different levels of the biological sample;

constructing a three-dimensional image indicating discrete signals at different levels of the three-dimensional image; and

counting the discrete signals at different levels of the three-dimensional image.

16. The method of claim 15, wherein the three-dimensional image is constructed by determining a location of a signal segment in the different levels of the biological sample, combining overlapping signal segments in contiguous levels into a single spot signal, and separating signal segments in non-contiguous levels into different spots.

17. The method of claim 16, wherein the location of signal segments is determined by the presence of an increase in brightness intensity that indicates an increase of signal as compared to a background signal.

18. The method of claim 17, wherein the probes display fluorescent signals, and the increase in brightness intensity is associated with an increase in fluorescence compared to the background signal.

19. The method of claim 15, wherein the signals comprise test signals from a test probe and reference signals from a reference probe.

20. The method of claim 19, wherein the test probe recognizes a gene of interest, and the reference probe recognizes a chromosomal locus having an expected quantity in the biological specimen.

21. The method of claim 20, further comprising determining a ratio between the test signals and the reference signals.

22. The method of claim 21, further comprising determining:

(a) whether there is an increase in an expected ratio between the test signal and the reference signal, indicating an amplification of the gene of interest; or

(b) whether there is a decrease in the expected ratio between the test signal and the reference signal, indicating relative loss of the gene of interest.

23. The method of claim 19, wherein the test probe is selected from the group consisting of probes that recognize genes implicated or suspected in the development or progression of a tumor.

24. The method of claim 15, wherein the biological samples are in a microarray.

25. The method of claim 24, wherein the microarray comprises a tissue microarray.

26. The method of claim 25, wherein the tissue microarray comprises tissue samples of a same tissue type taken from a plurality of donor specimens.

27. The method of claim 15, wherein the plurality of images consists of between eight and thirty two images at different levels of the biological sample.

28. The method of claim 15, further comprising:

avoiding counting discrete signals having intensities exceeding a threshold intensity.

29. The method of claim 15, further comprising:

avoiding counting discrete signals having a combined intensity and area exceeding a threshold value.

30. The method of claim 15, further comprising:

avoiding counting discrete signals related to autofluorescent material.

31. The method of claim 15, further comprising:

depicting a two-dimensional image representing the three-dimensional image for consideration by a user.

32. The method of claim 31, further comprising:

emphasizing discrete signals related to autofluorescent material in the two-dimensional image.

33. The method of claim 15, further comprising:

identifying a set of one or more discrete signals as a cluster; and

counting the cluster as a number of discrete signals greater than the number of discrete signals in the set.

34. The method of claim 33 wherein the cluster is counted as a number of discrete signals indicated by applying a mapping to the number of discrete signals in the set.

35. The method of claim 33 wherein the cluster is counted as a number of discrete signals indicated by a function calibrated via manual counting of spots in a plurality of images.

36. The method of claim 33 wherein the cluster is counted as a number of discrete signals indicated by a gain factor applied to the number of discrete signals in the set.

37. The method of claim 15 wherein the plurality of images are a set of images taken during a first analysis of a first color channel, and a second set of images are taken of the biological sample for a second color channel, the method further comprising:

avoiding counting discrete signals appearing at a same location in the set of images for the first color channel and the set of images in the second color channel.

38. The method of claim 15 wherein the plurality of images are a set of images taken for a test probe, and a second set of images are taken of the biological sample for a reference probe, the method further comprising:

avoiding counting discrete signals appearing at a same location in the set of images for the test probe and the set of images for the reference probe.

39. The method of claim 15 further comprising:

receiving a directive from a user indicating counting is to be avoided for a specified portion of the biological sample; and

responsive to the directive, avoiding counting discrete signals for the specified portion of the biological sample.

40. The method of claim 15 further comprising:

receiving a directive from a user indicating counting is to be performed separately for a specified portion of the biological sample; and

responsive to the directive, separately counting discrete signals for the specified portion of the biological sample.

41. A high-throughput method of counting fluorescent in situ hybridization signals, comprising:

providing an array of biological samples;

hybridizing the biological samples with a fluorescent test probe that hybridizes to a gene of interest in the biological samples and with a fluorescent reference probe that hybridizes to a chromosomal reference locus in the biological samples;

obtaining images by confocal microscopy of contiguous sections at different depths of a plurality of the biological samples;

detecting fluorescent signal segments from the contiguous sections, and segmenting contiguous signals in different contiguous sections into a corresponding single spot signal; and

projecting each spot signal into a two-dimensional plane, and counting the spots.

42. The method of claim 41, wherein the fluorescent spot signals from the test probe are obtained and detected separately from the fluorescent spot signals from the reference probe.

43. The method of claim 42, wherein the fluorescent spot signals from the test probe are obtained and detected simultaneously with the fluorescent spot signals from the reference probe.

44. The method of claim 41, wherein the array comprises an array of at least 6 different specimens.

45. The method of claim 44, wherein fluorescent in situ hybridization signals are counted in at least 50 different specimens.

46. The method of claim 41, wherein the method is a computer implemented system.

47. The method of claim 41, wherein providing an array of biological samples comprises providing an array of biological specimens:

obtaining a plurality of donor specimens;

placing each donor specimen in an assigned location in a recipient array; and

obtaining a plurality of copies of the recipient array in a manner that each copy contains a plurality of donor specimens that maintain their assigned locations.

48. A method for counting signals from in situ hybridization of probes in biological tissue, the method comprising:

obtaining digital images of different levels of the biological tissue with a confocal microscope;

detecting spot signals at the different levels and separating overlying signals from one another.

49. The method of claim 48, wherein the spot signals are detected at different levels with a computer implemented system that detects fluorescent signals as signal segments from the different levels, wherein the levels are contiguous optical sections, and the system segments contiguous signal segments from different vertical levels into one spot signal, while separating non-contiguous signal segments from different levels into different spot signals.

50. The method of claim 49, wherein the computer implemented system projects the spot signals into a single level and counts the spot signals.

51. The method of claim 49, wherein the computer implemented system detects fluorescent signals as spot signals by

(a) performing a morphological transform to digital images of the different levels to obtain digital representations of brightness intensity that indicate a signal segment;

(b) eliminating signal segments below a threshold;

(c) grouping contiguous signal segments into a single spot.

52. The method of claim 51, further comprising:

avoiding signal segments having an intensity above a threshold.

53. The method of claim 51, further comprising:

identifying clusters within the signal segments; and

counting the clusters each as a plurality of spots.

54. A device for counting signals from in situ hybridization of probes in biological tissue, the device comprising:

a confocal microscope;

a digital camera positioned to obtain digital optical sections of different levels of the biological tissue; and

a computer implemented system that detects brightness signals at the different levels and separates overlapping brightness signals from one another.

55. The device of claim 54, wherein the computer implemented system detects fluorescent signals as signal segments from the different levels, wherein the levels are contiguous, and the system groups contiguous signal segments from different levels into one spot signal, while separating non-contiguous signal segments from different levels into different spot signals.

56. The device of claim 55, wherein the computer implemented system counts the spot signals.

57. The device of claim 55, wherein the computer implemented system detects fluorescent signals as spot signals by

(a) performing a morphological transform to digital images of the different levels to obtain a digital representation of brightness intensity that indicate a signal segment;

(b) eliminating signal segments below a threshold;

(c) segmenting contiguous signal segments; and

(d) grouping vertical contiguous signal segments into a single spot signal at an assigned level associated with a greatest signal segment intensity.

58. A computer-readable medium comprising computer-executable instructions for performing the following:

within a stack of image slices generated from a plurality of confocal microscopic observations of a FISH experiment as a plurality of depths along a z-axis, identifying possible fluorescent image components;

projecting the possible fluorescent image components within the image slices onto a projection image;

discarding insignificant contiguous possible fluorescent image components in the slices;

for each contiguous region in the projection image, grouping regions of possible florescent image components associated with the contiguous region in the projection image into spot candidates;

applying a filter to the spot candidates; and

counting the remaining spot candidates as spots.

59. The computer-readable medium of claim 58 wherein insignificant contiguous possible fluorescent image components are determined by comparing a size of a contiguous possible fluorescent image component with a threshold.

60. A computer-generated user interface for presenting results of microscopic observation of biological tissue subjected to a FISH experiment, the user interface comprising:

a scatter plot of sets of image components designated as spot candidates for the FISH experiment;

wherein the scatter plot comprises points indicating a size and intensity of spot candidates.

61. The computer-generated user interface of claim 60 wherein at least one point is operable to receive a user interface activation to navigate to a user interface display of information for a spot candidate associated with the point.

62. The computer-generated user interface of claim 60 wherein at least one point is operable to receive a user interface activation to navigate to a user interface display of a three-dimensional depiction of a spot candidate associated with the point.

63. The computer-generated user interface of claim 60 further comprising:

a display image depicting a view of the tissue subjected to the FISH experiment and a depiction of least one candidate spot thereon;

wherein the depiction of the candidate spot is operable to receive a user interface activation to designate the spot as a minimal intensity spot; and

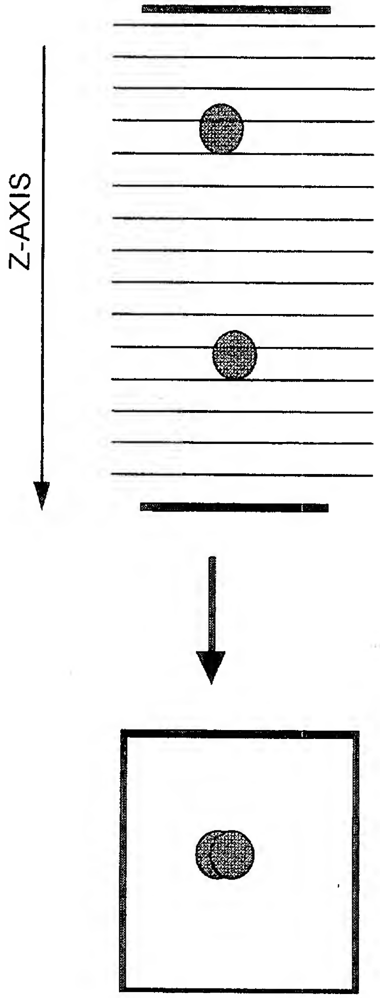
wherein candidate spots designated as minimal intensity spots are visually emphasized when presenting the scatter plot.

SIGNAL COUNTING FOR IN SITU HYBRIDIZATION

ABSTRACT

Fluorescently tagged nucleic acid probe signals are counted in biological specimens by determining a ratio of signals from a test probe to signals of a reference probe. Probe signals need not be counted with reference to cells, nuclei, or nuclear contours. Gene amplification or deletion can thus be detected by analyzing the ratio. Successive image slices are obtained by confocal microscopy, and the images are digitized. The digital images are transformed and analyzed to combine contiguous fluorescent signal segments in successive optical sections to identify discrete probe signals, or spots. Spots overlapping in the axial and transverse dimensions of a three-dimensional representation of the biological specimens can be distinguished. A graphical user interface presents various features for consideration by a user, who can provide guidance to a computer system counting the spots. Various features directed to identifying spot clusters and autofluorescent material can increase accuracy of spot counting.

FIG. 1A



STACK OF CONFOCAL IMAGES

OUT-OF-FOCUS DOTS ARE VISIBLE IN NON-CONFOCAL IMAGES, RESULTING IN LARGE, BLURRED SIGNALS

NON-CONFOCAL IMAGE

FIG. 1B

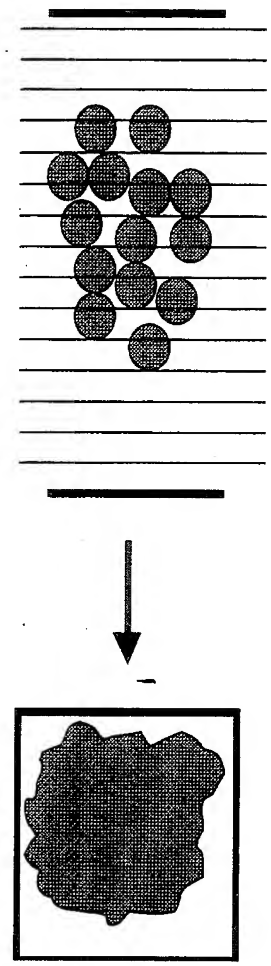


FIG. 2A

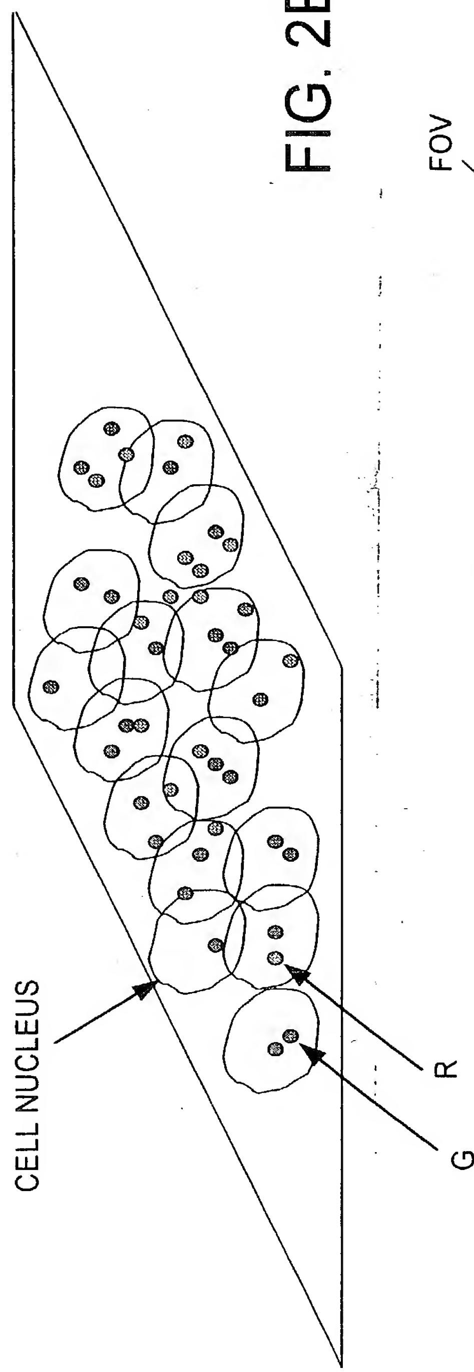


FIG. 2B

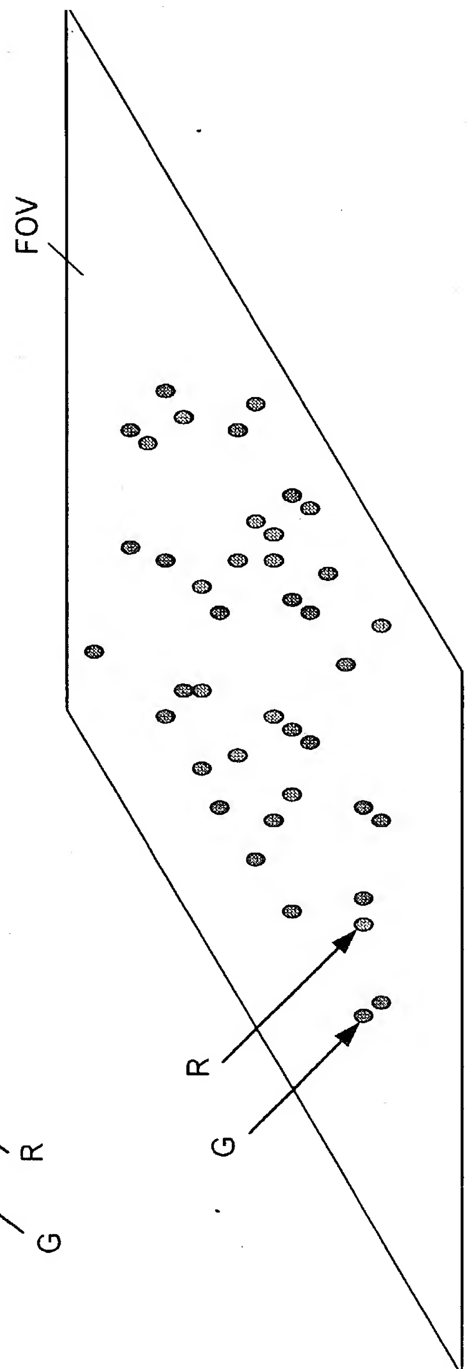
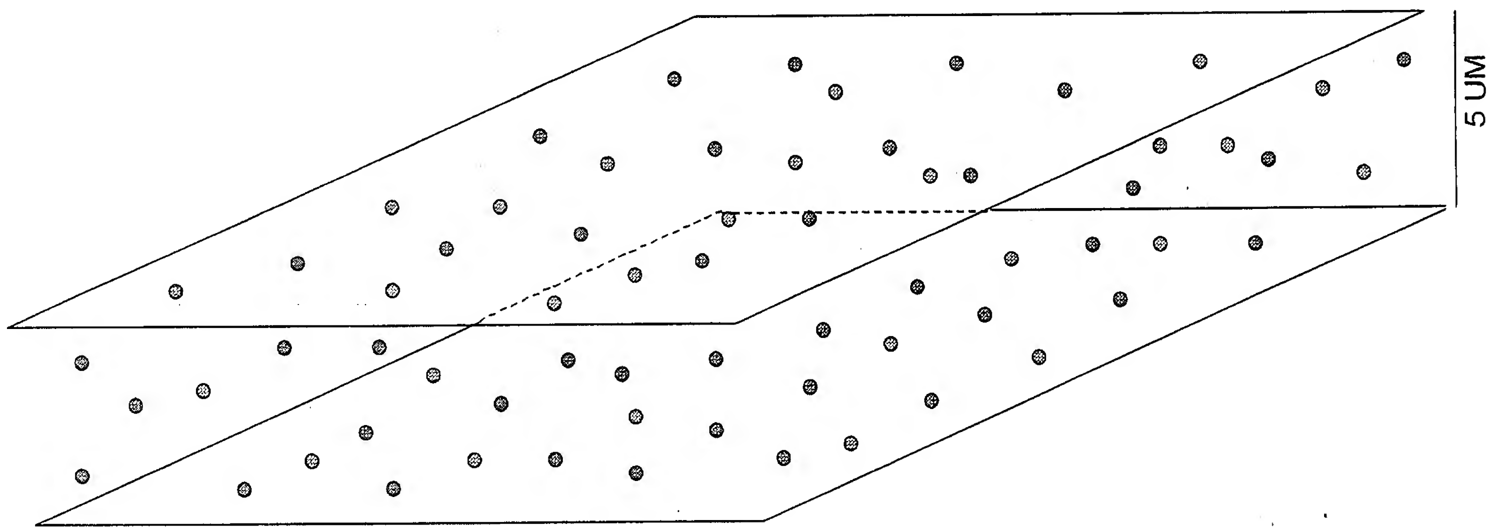


FIG. 2C

REGIONS OF INTEREST CAN DEFINE AN AREA WHICH IS SELECTED FOR A SPECIFIC ANALYSIS. IN THIS EXAMPLE OF A PROSTATE SPECIMEN ON A TISSUE MICROARRAY (H&E STAINING, X100), ONLY THE AREA CONTAINING CANCER TISSUE WILL BE SELECTED FOR FISH ANALYSIS. STROMAL TISSUE AND NORMAL GLANDS WILL BE EXCLUDED.

FIG. 3



RED/GREEN PROBE SIGNAL RATIOS ARE FORMED BY DOT IDENTIFICATION AND COUNTING FROM BOTH COLOR CHANNELS FROM CONSECUTIVE SERIES OF CONFOCAL MICROSCOPE IMAGES

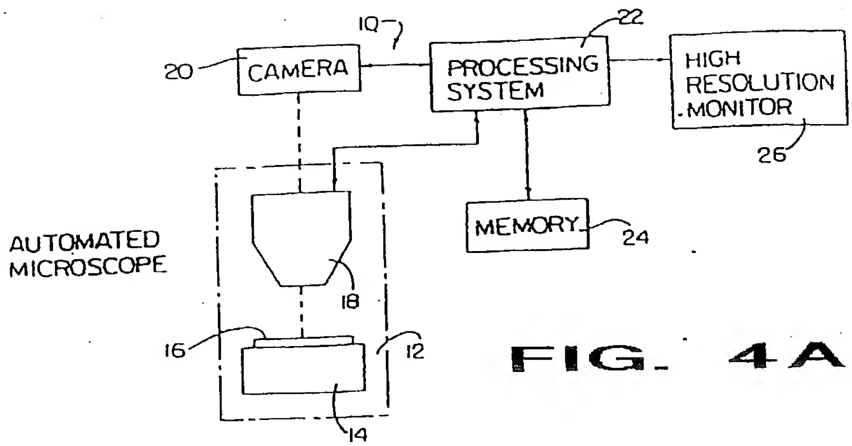

10
20
CAMERA
22
PROCESSING SYSTEM
HIGH RESOLUTION MONITOR
26
MEMORY
24
AUTOMATED MICROSCOPE
18
16
12
14

FIG. 4A

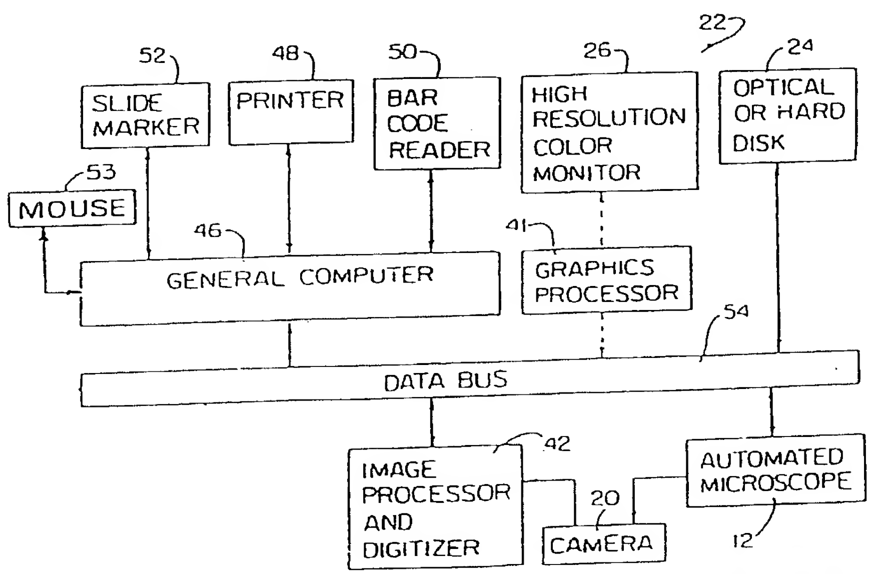
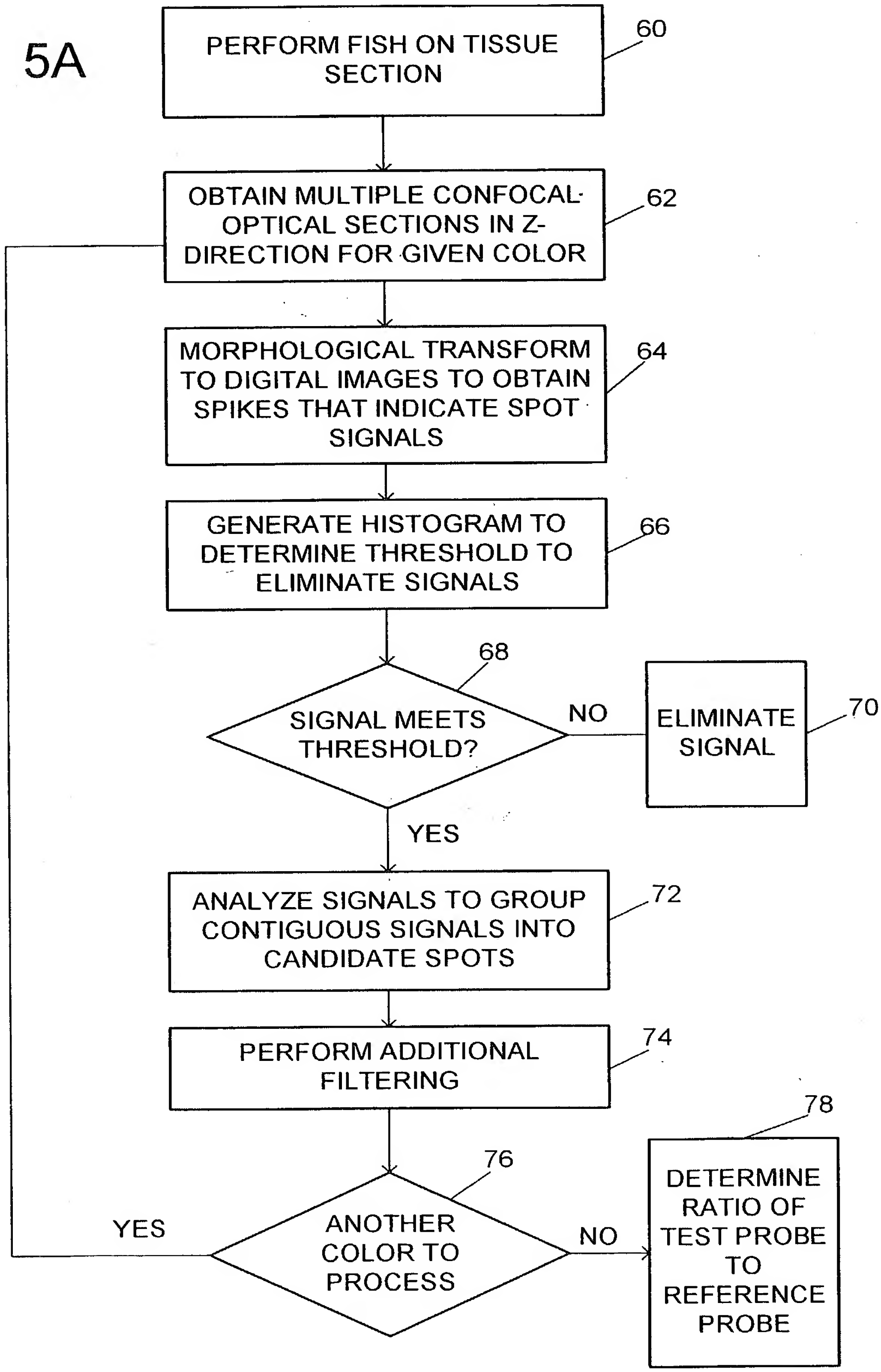

22
52
SLIDE MARKER
48
PRINTER
50
BAR CODE READER
26
HIGH RESOLUTION COLOR MONITOR
24
OPTICAL OR HARD DISK
53
MOUSE
46
GENERAL COMPUTER
41
GRAPHICS PROCESSOR
54
DATA BUS
42
IMAGE PROCESSOR AND DIGITIZER
20
CAMERA
AUTOMATED MICROSCOPE
12

FIG. 4B

FIG. 5A


PERFORM FISH ON TISSUE SECTION
60
OBTAIN MULTIPLE CONFOCAL OPTICAL SECTIONS IN Z-DIRECTION FOR GIVEN COLOR
62
MORPHOLOGICAL TRANSFORM TO DIGITAL IMAGES TO OBTAIN SPIKES THAT INDICATE SPOT SIGNALS
64
GENERATE HISTOGRAM TO DETERMINE THRESHOLD TO ELIMINATE SIGNALS
66
68
SIGNAL MEETS THRESHOLD?
NO
ELIMINATE SIGNAL
70
YES
ANALYZE SIGNALS TO GROUP CONTIGUOUS SIGNALS INTO CANDIDATE SPOTS
72
PERFORM ADDITIONAL FILTERING
74
78
76
YES
ANOTHER COLOR TO PROCESS
NO
DETERMINE RATIO OF TEST PROBE TO REFERENCE PROBE

10/088269

FIG. 5B

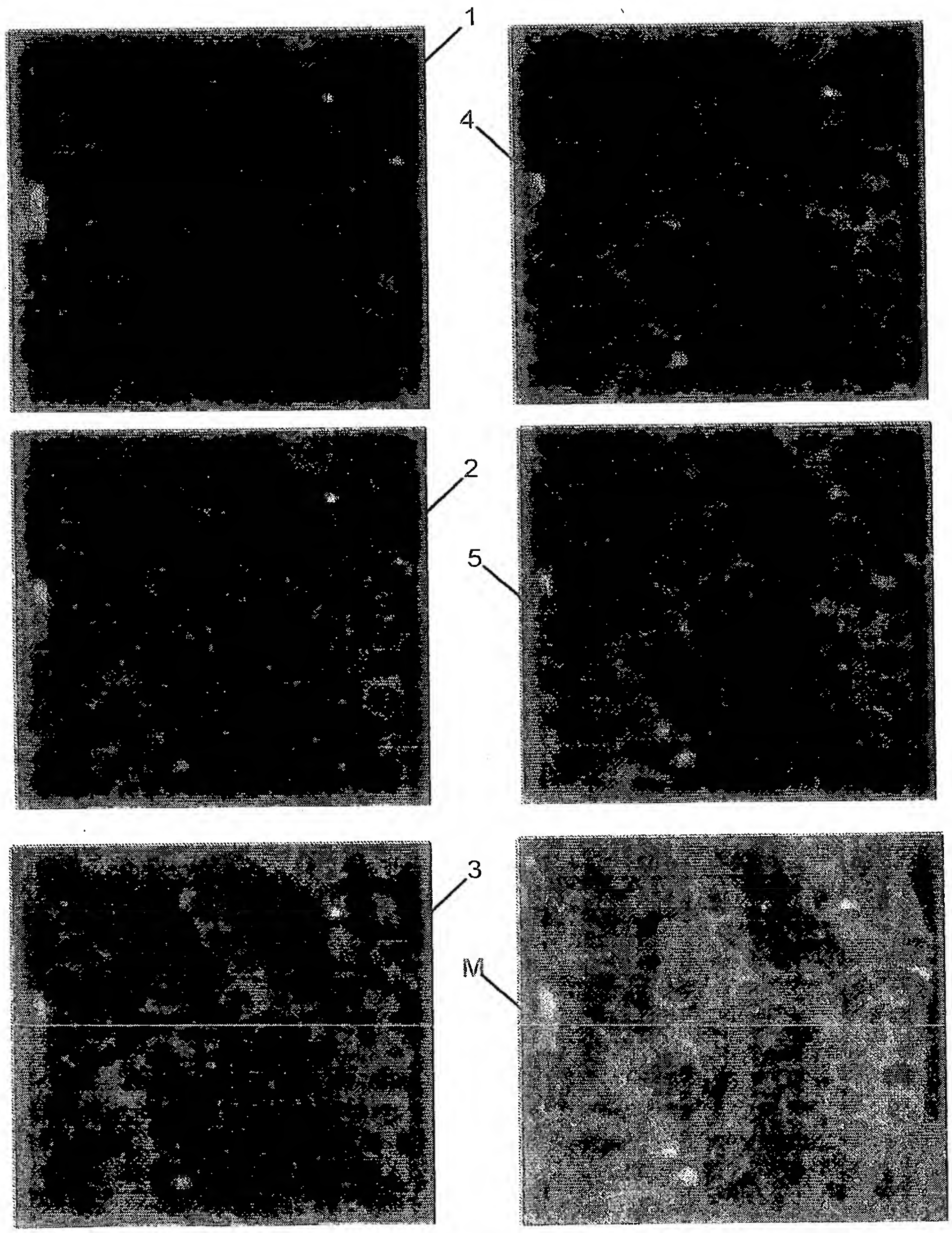
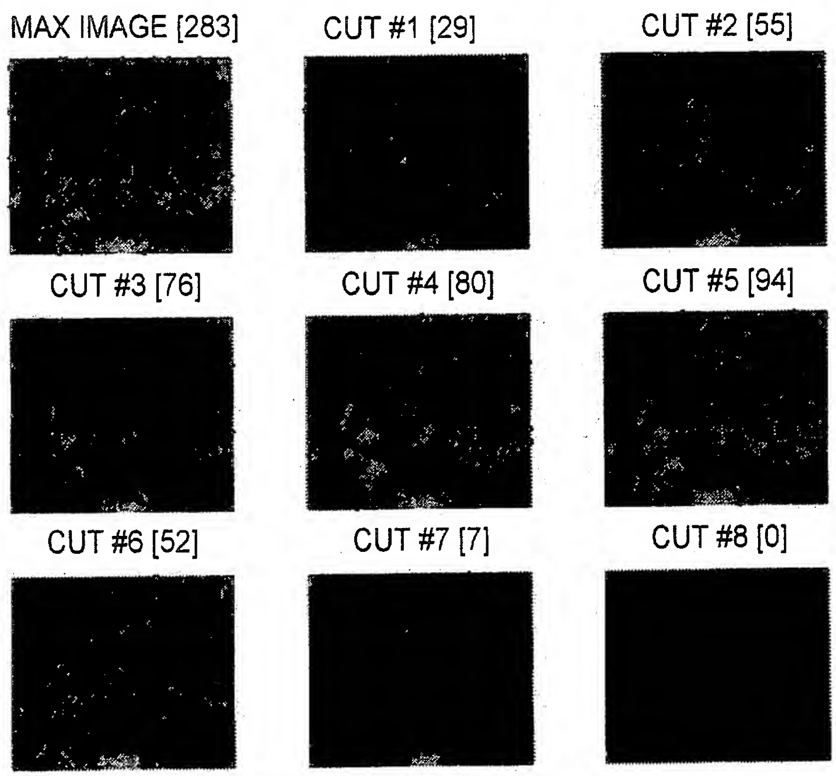


FIG. 6A

AN EXAMPLE OF AUTOMATED FISH-SIGNAL COUNTING
FROM A SERIES OF CONFOCAL IMAGES



STACK 7 FITC (T=21)

AUTOMATED SPOT COUNTING RESULTS:

TOTAL NUMBER OF SIGNALS COUNTED: 393
- UNIQUE SIGNALS *: 283

RESULTS FROM MANUAL COUNTING

TOTAL NUMBER OF SIGNALS COUNTED: 391
- UNIQUE SIGNALS : IMPOSSIBLE TO DETERMINE

* FISH SPOT MAY APPEAR IN SEVERAL FOCAL LEVELS; THEREFORE, THE NUMBER OF UNIQUE SIGNALS IS LOWER THAN THE NUMBER OF ALL SIGNALS

FIG. 6B

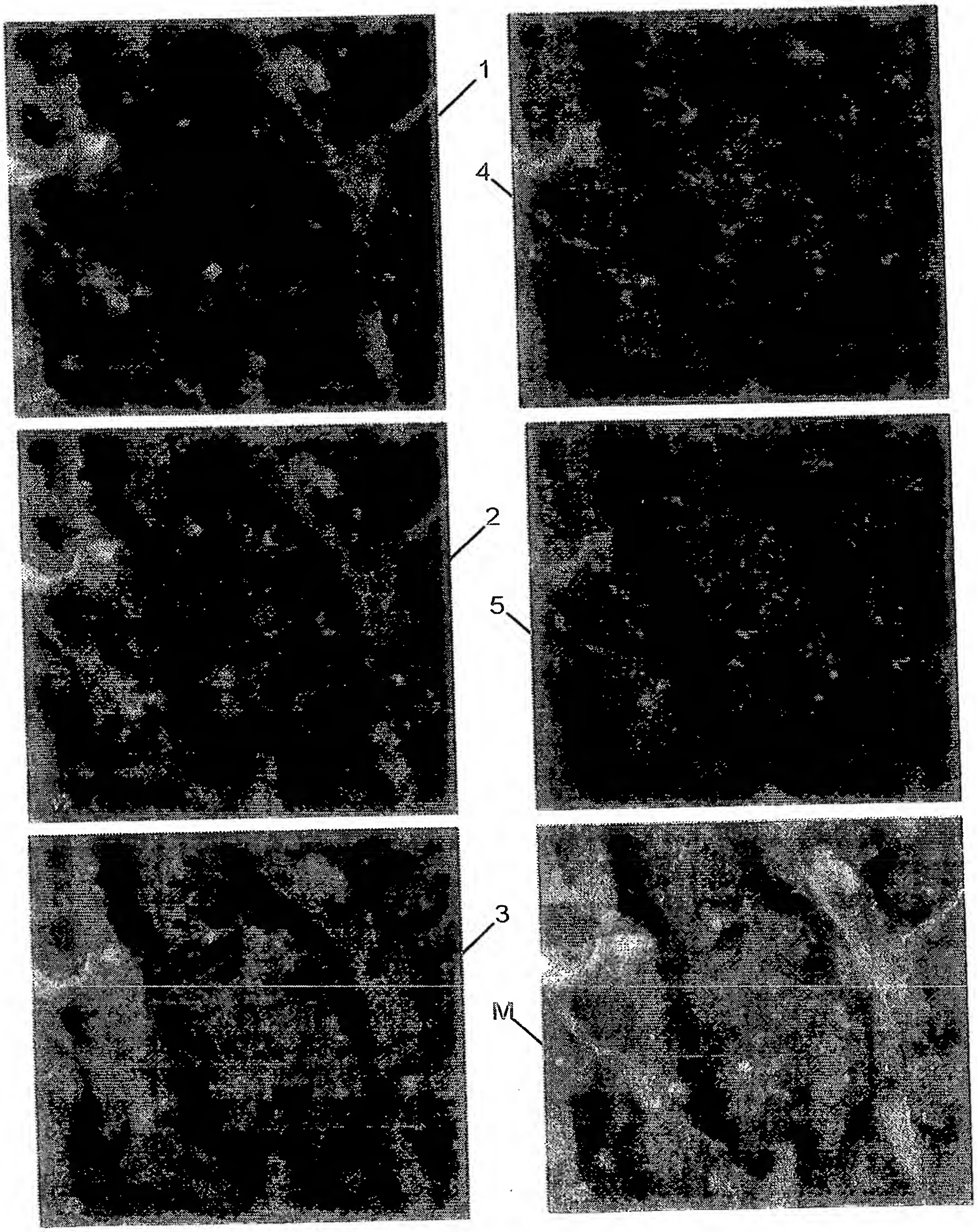


FIG. 7

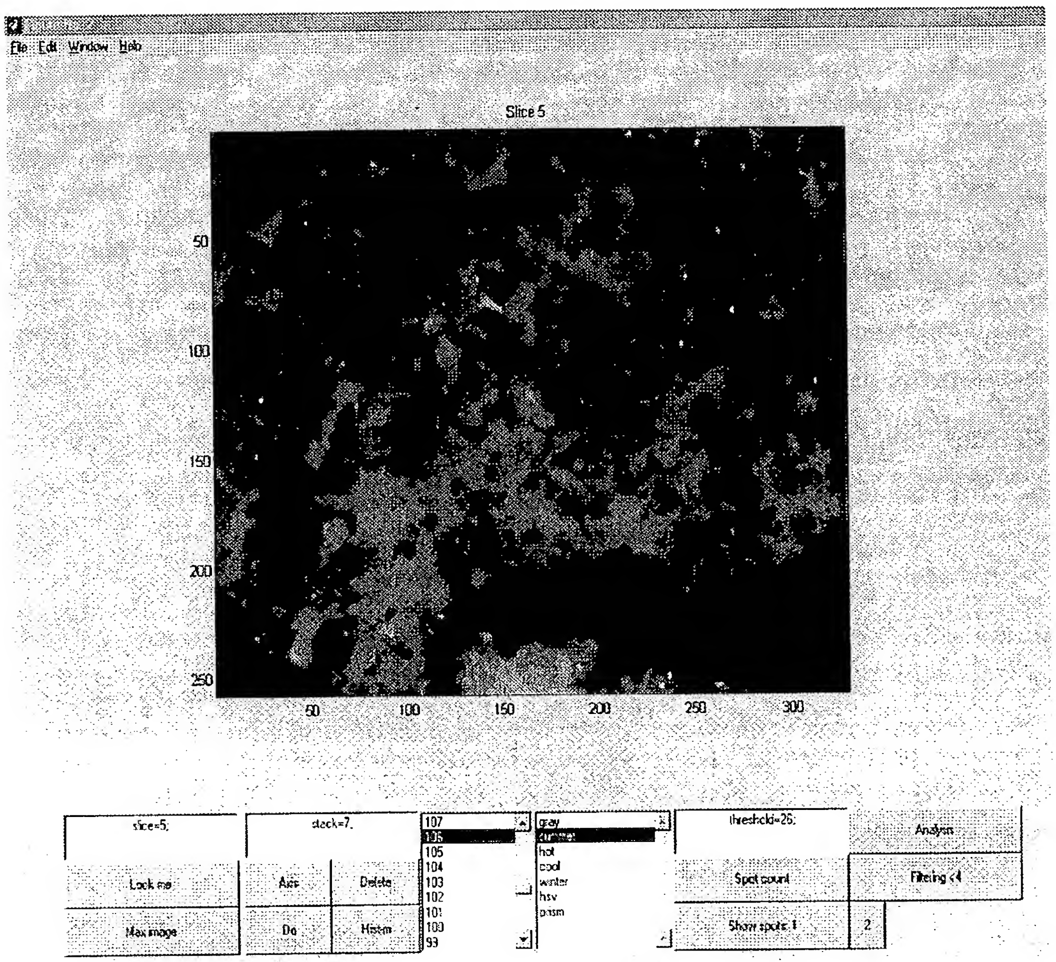

File Edit Window Help
Slice 5
50
100
150
200
250
50
100
150
200
250
300
slice=5;
stack=7;
Lock me
Max image
Add
Delete
Do
Histm
107
106
105
104
103
102
101
100
99
gray
summer
hot
cool
winter
hsv
prism
threshold=26;
Analysis
Spot count
Filtering <4
Show spots: 1
2

FIG. 8

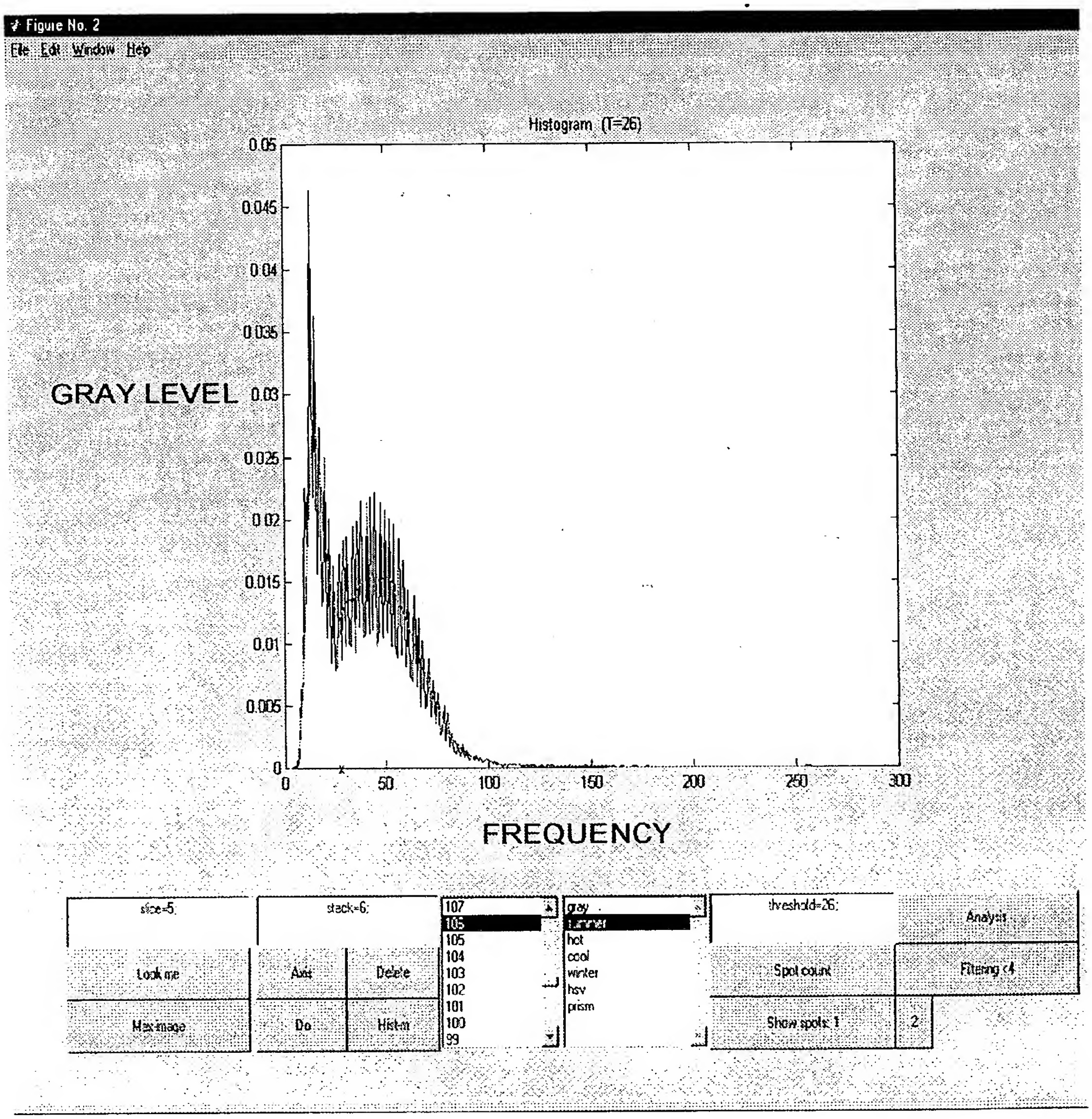
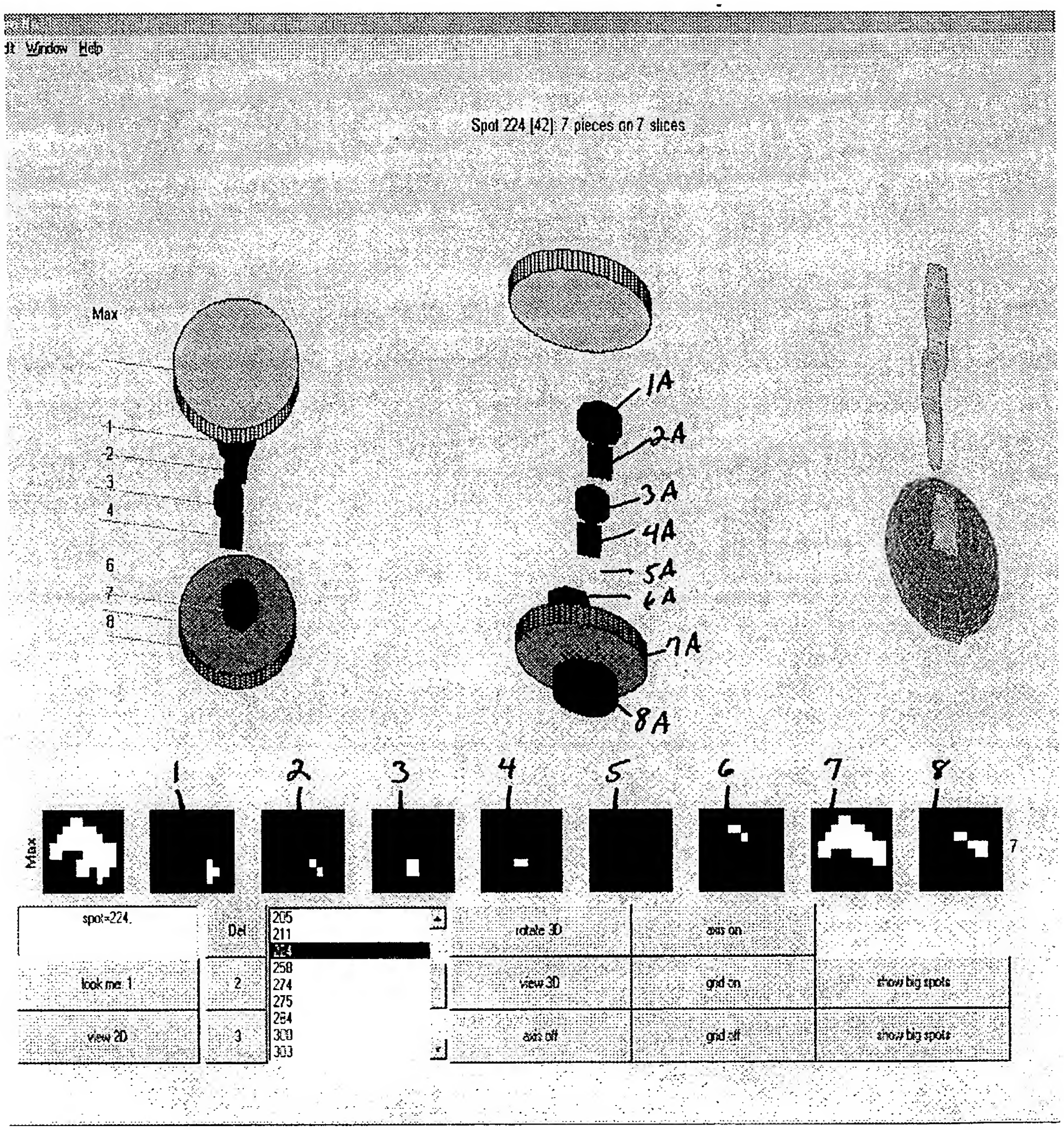


Figure No. 2
File Edit Window Help
Histogram (T=26)
GRAY LEVEL
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0
50
100
150
200
250
300
FREQUENCY
slice=5;
stack=6;
107
106
105
104
103
102
101
100
99
gray
summer
hot
cool
winter
hsv
prism
threshold=26;
Analysis
Look me
Axe
Delete
Spot count
Filtering c4
Max-image
Do
Hist-m
Show spots: 1
2

FIG. 9A



Window Help
Spot 224 [42]: 7 pieces on 7 slices
Max
1
2
3
4
6
7
8
1A
2A
3A
4A
5A
6A
7A
8A
Max
1
2
3
4
5
6
7
8
7
spot=224.
Del
look me: 1
2
view 2D
3
205
211
224
258
274
275
284
300
303
rotate 3D
view 3D
axis off
axis on
grid on
grid off
show big spots
show big spots

FIG. 9B

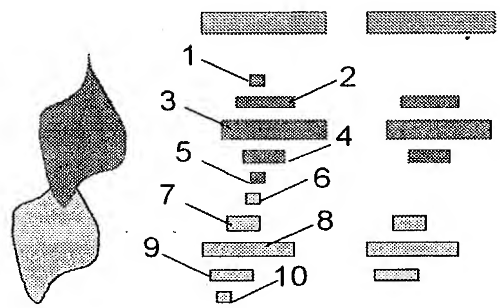


FIG. 9C

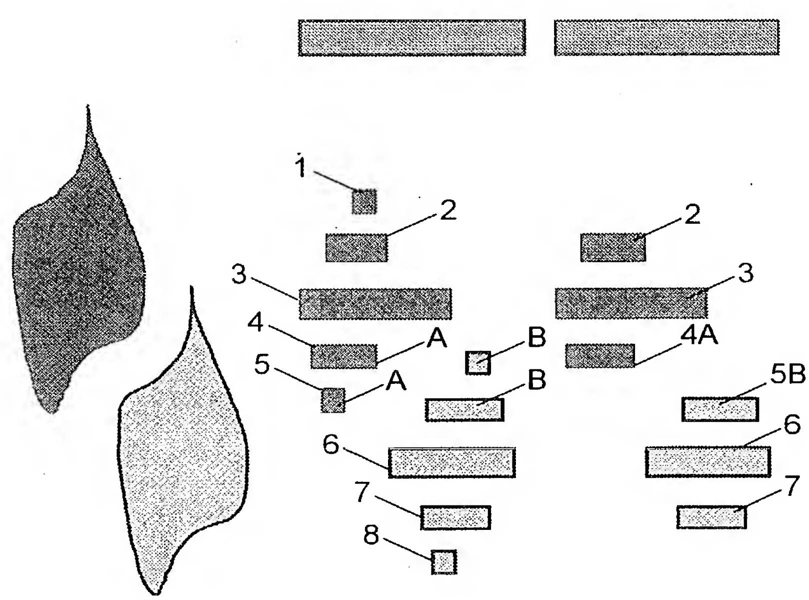
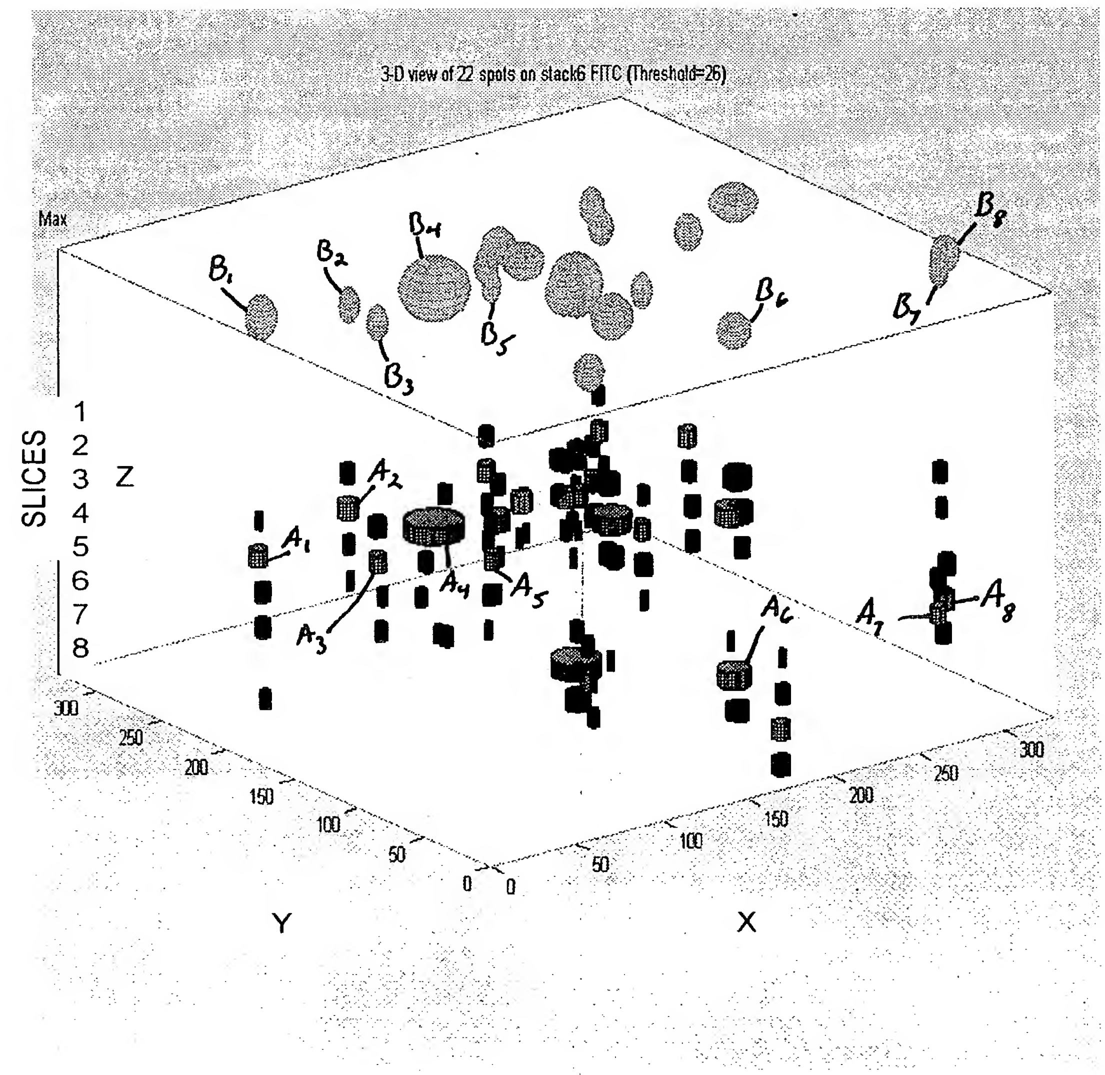
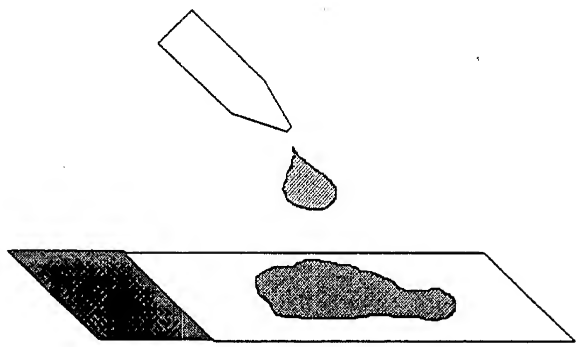


FIG. 10



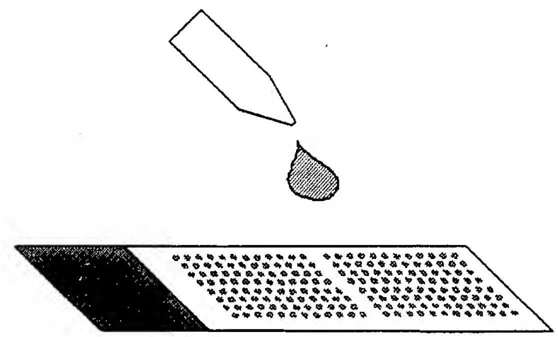
ROTATE 3D	VIEW 3	VIEW 2	GRID ON	GRID OFF	GO BACK

FIG. 11



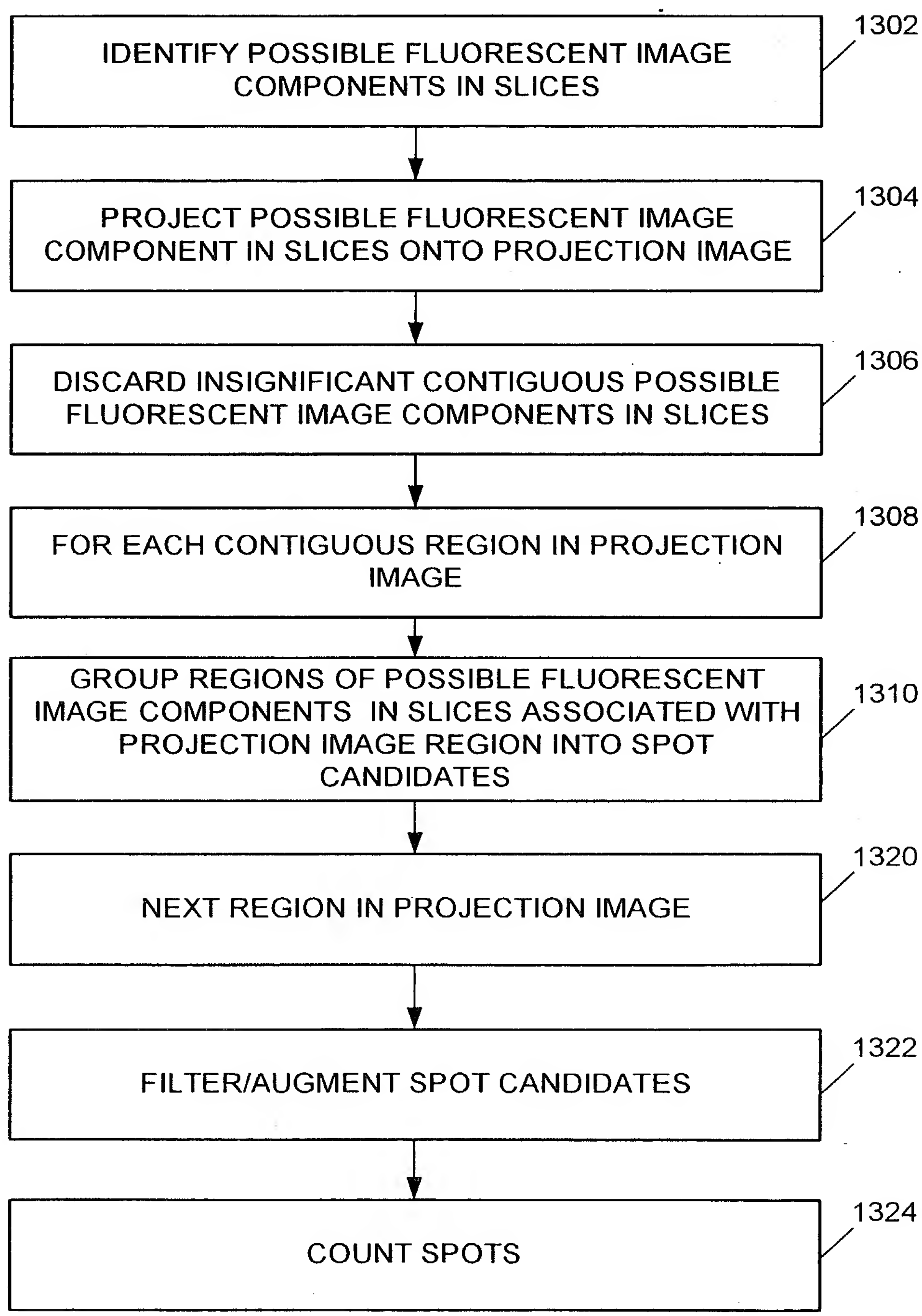
1 TUMOR

FIG. 12



1000 TUMORS

FIG. 13


IDENTIFY POSSIBLE FLUORESCENT IMAGE COMPONENTS IN SLICES
1302
PROJECT POSSIBLE FLUORESCENT IMAGE COMPONENT IN SLICES ONTO PROJECTION IMAGE
1304
DISCARD INSIGNIFICANT CONTIGUOUS POSSIBLE FLUORESCENT IMAGE COMPONENTS IN SLICES
1306
FOR EACH CONTIGUOUS REGION IN PROJECTION IMAGE
1308
GROUP REGIONS OF POSSIBLE FLUORESCENT IMAGE COMPONENTS IN SLICES ASSOCIATED WITH PROJECTION IMAGE REGION INTO SPOT CANDIDATES
1310
NEXT REGION IN PROJECTION IMAGE
1320
FILTER/AUGMENT SPOT CANDIDATES
1322
COUNT SPOTS
1324

10/088269

FIG. 14

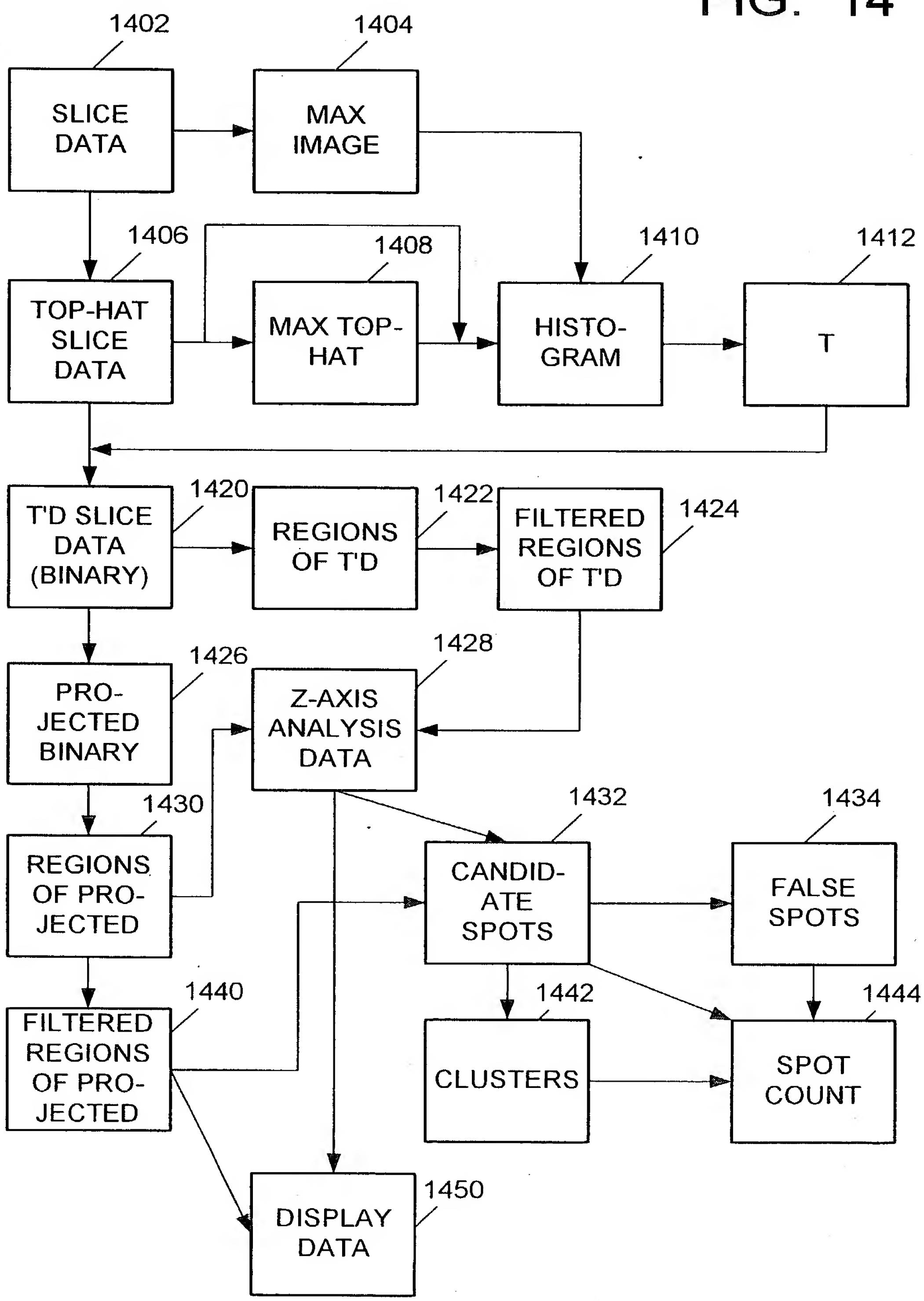


FIG. 15A

FIG. 15B

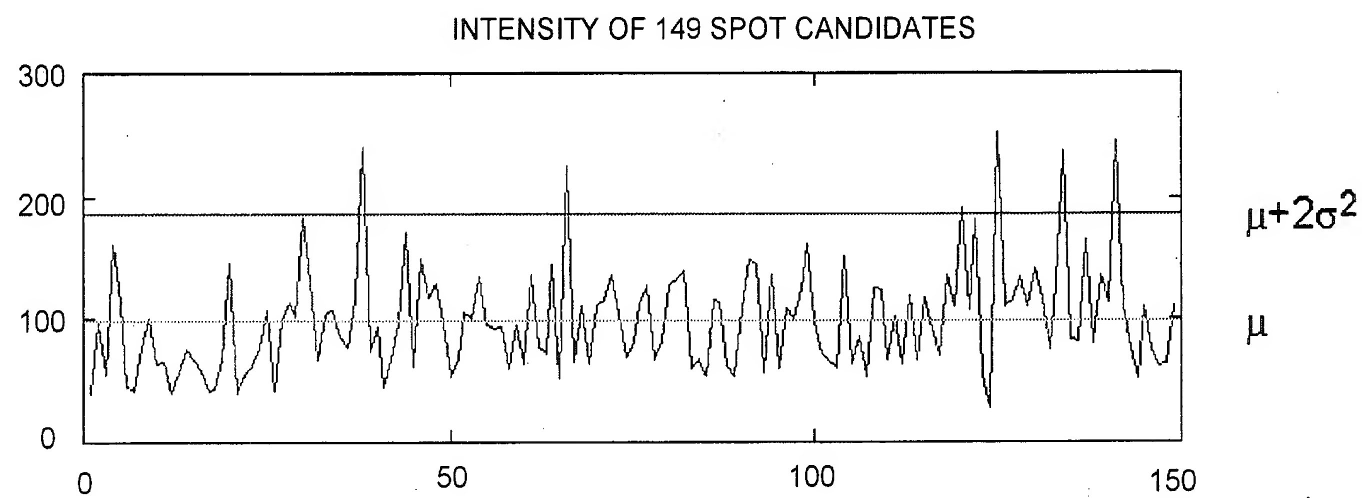

INTENSITY OF 149 SPOT CANDIDATES
300
200
100
0
0
50
100
150
μ+2σ2
μ

FIG. 15C

INTENSITY COMBINED WITH AREA FOR 149 SPOT CANDIDATES

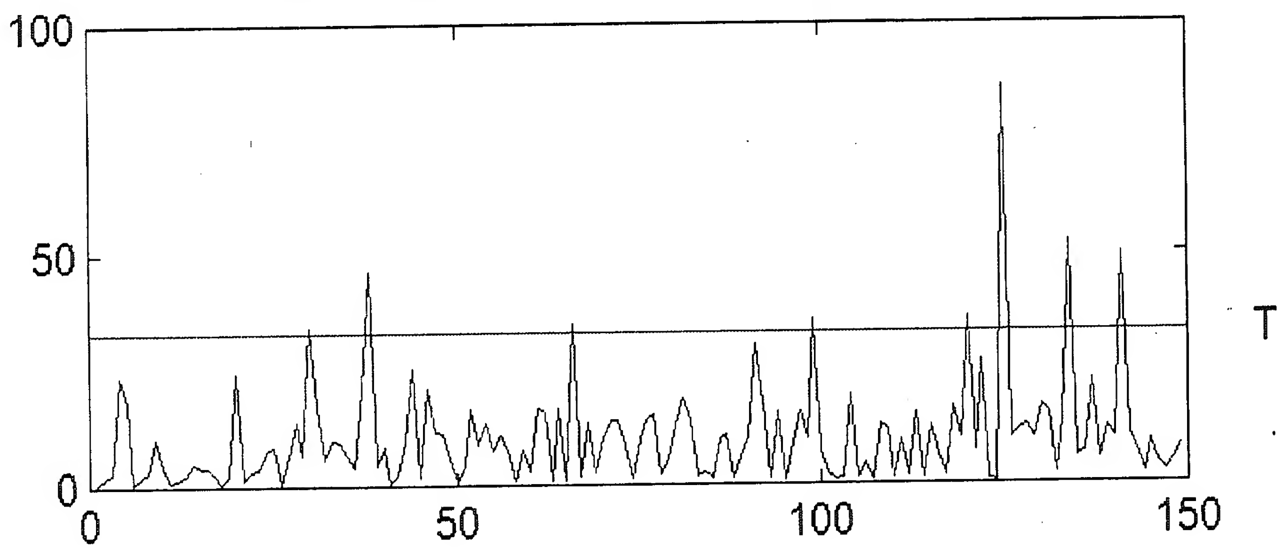
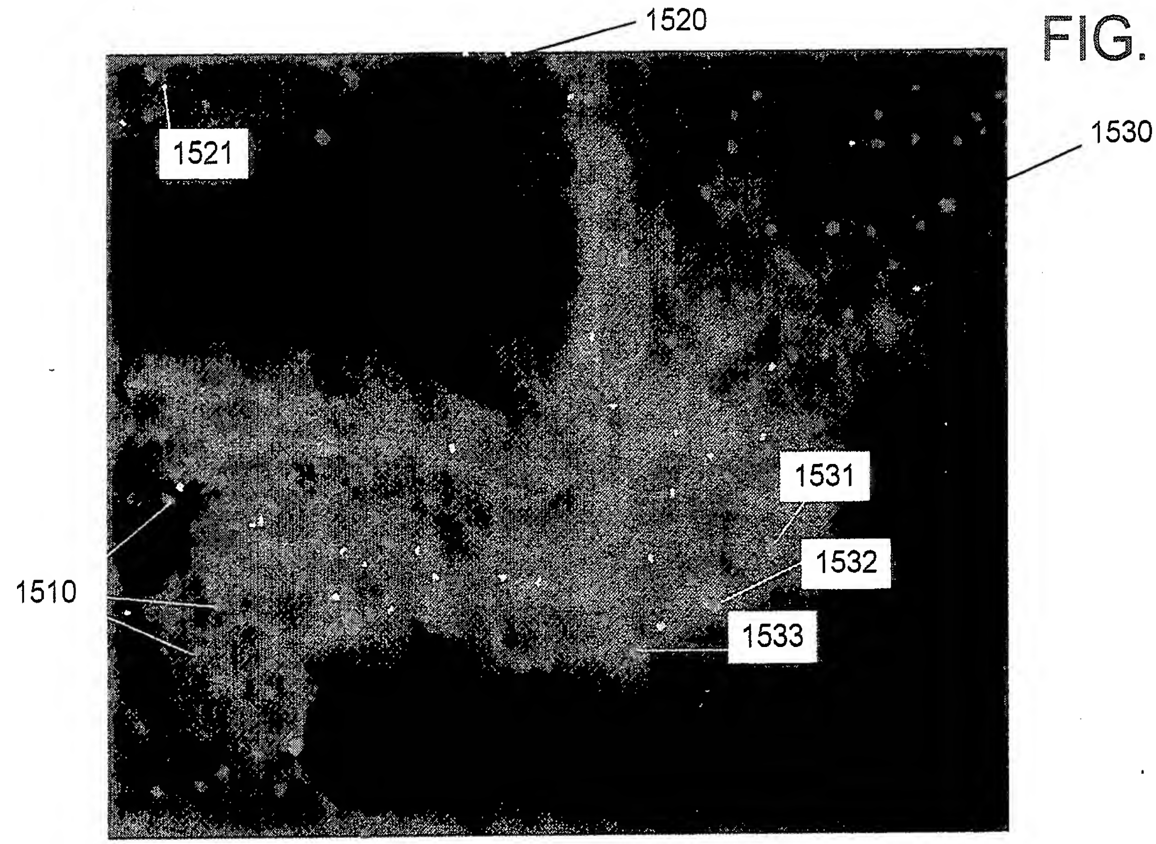


FIG. 15D



- true FISH signals
- small autofluorescent tissue particles
- large autofluorescent tissue particles

FIG. 16A

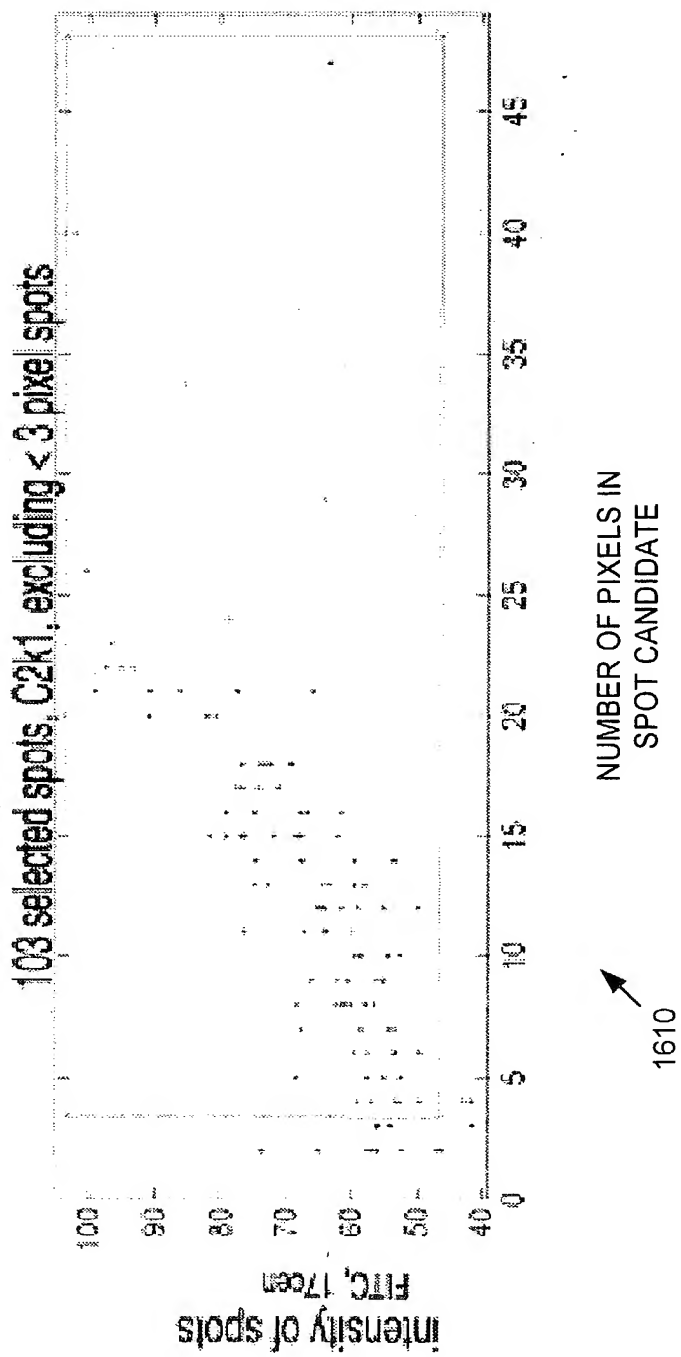

103 selected spots, C2K1, excluding < 3 pixel spots
intensity of spots
FITC, 17cem
100
90
80
70
60
50
40
0
5
10
15
20
25
30
35
40
45
NUMBER OF PIXELS IN SPOT CANDIDATE
1610

FIG. 16B

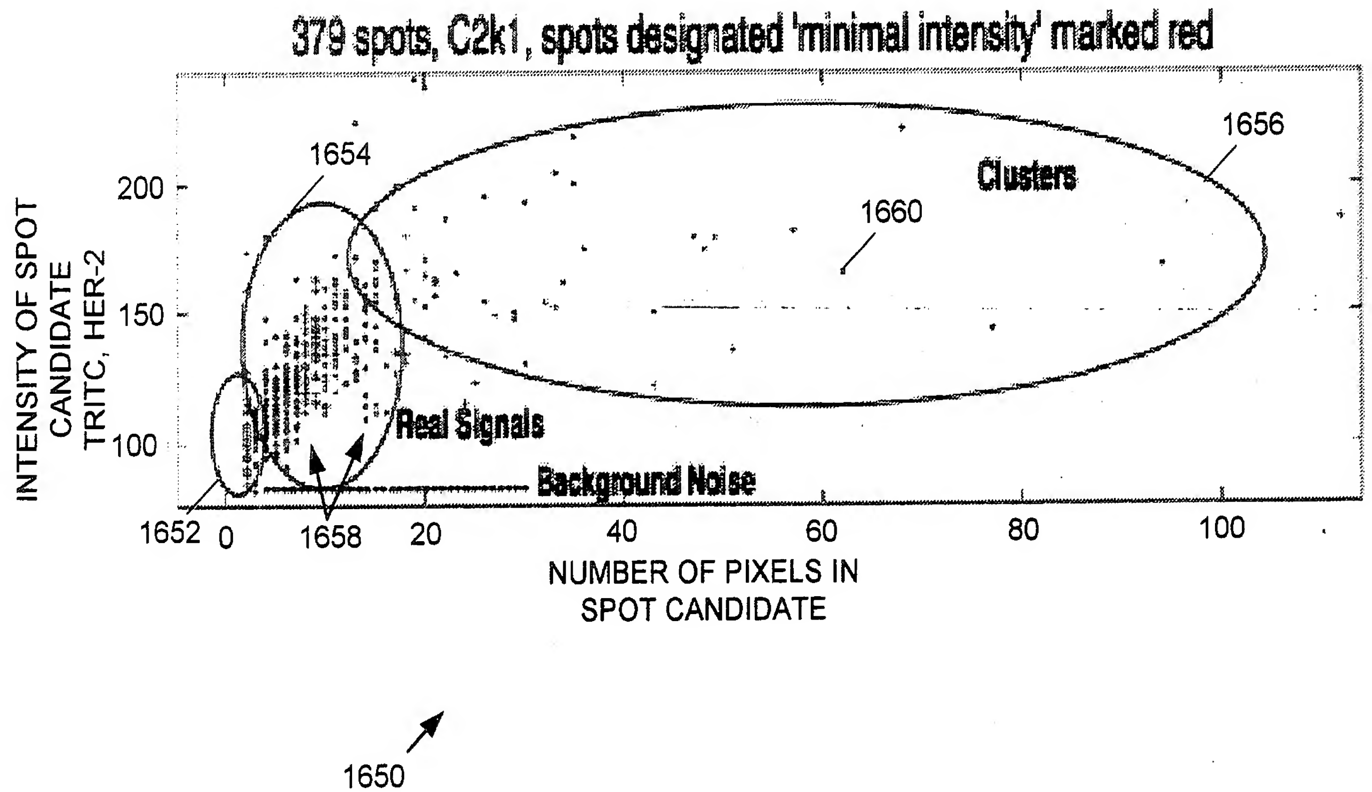

379 spots, C2k1, spots designated 'minimal intensity' marked red
INTENSITY OF SPOT CANDIDATE TRITC, HER-2
200
150
100
0
20
40
60
80
100
NUMBER OF PIXELS IN SPOT CANDIDATE
Clusters
Real Signals
Background Noise
1652
1654
1656
1658
1660
1650

FIG. 17

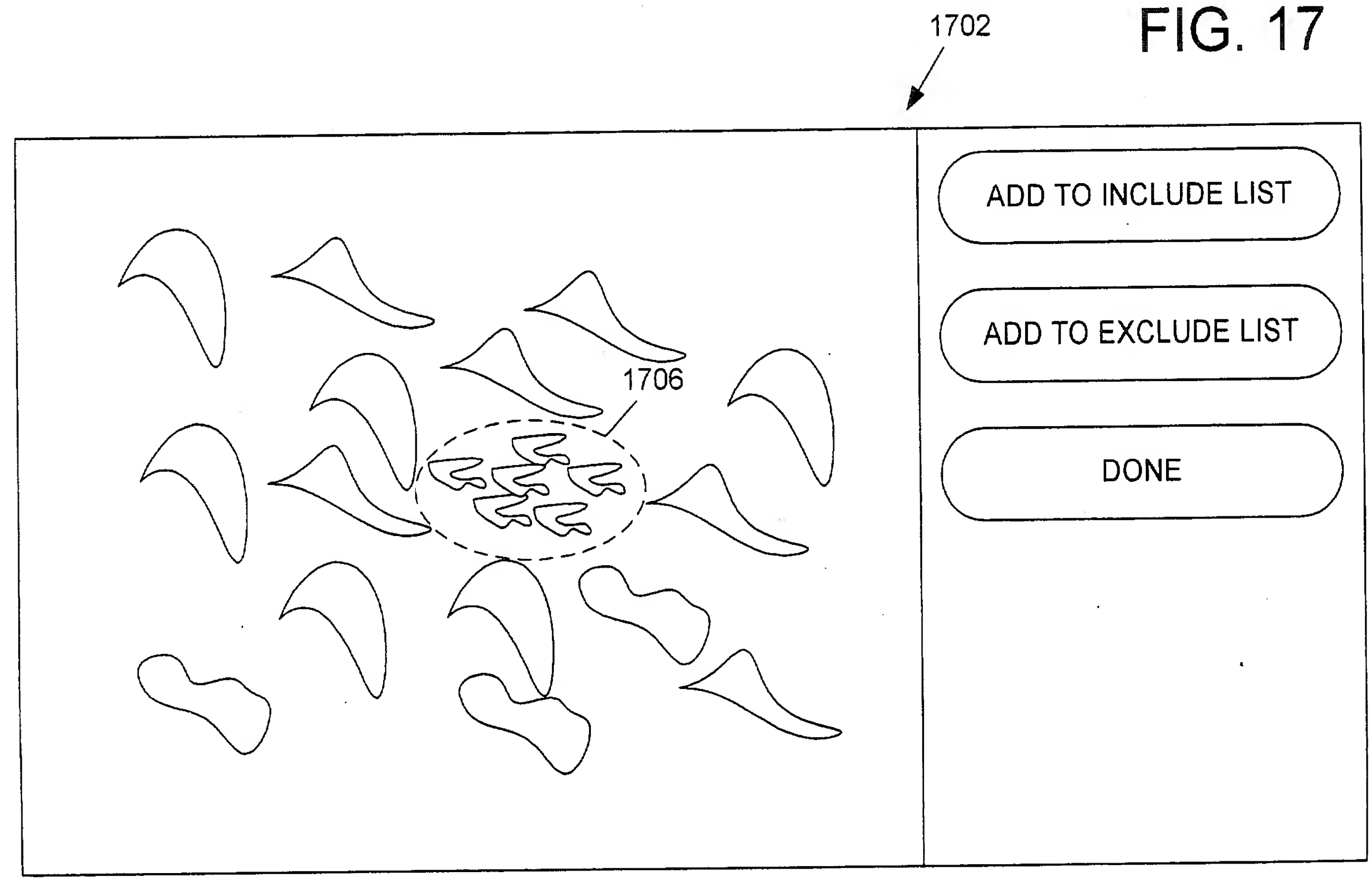
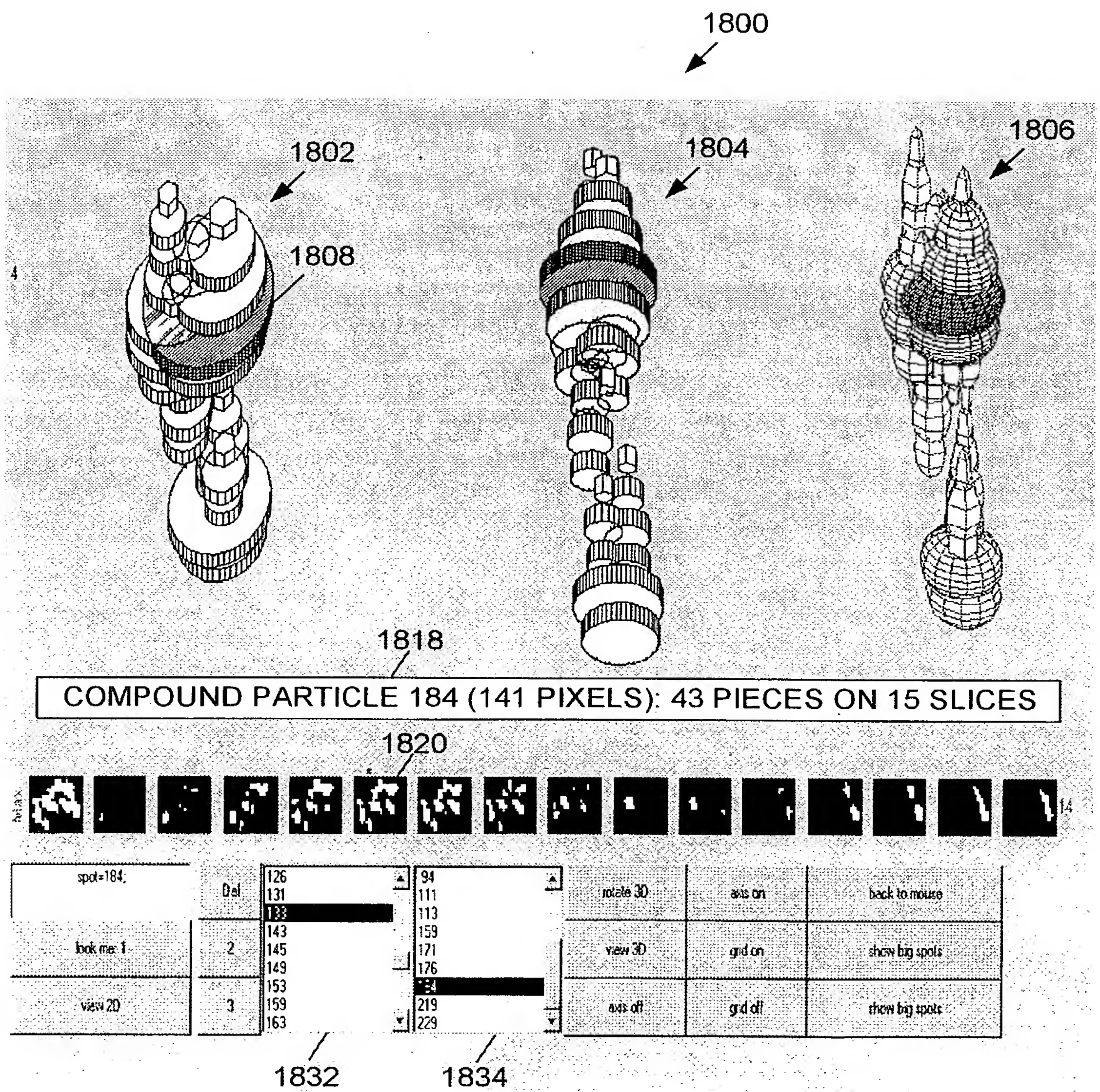


FIG. 18


1800
1802
1804
1806
1808
1818
COMPOUND PARTICLE 184 (141 PIXELS): 43 PIECES ON 15 SLICES
1820
spot=184;
look me: 1
view 2D
Del
2
3
126
131
133
143
145
149
153
159
163
94
111
113
159
171
176
219
229
rotate 3D
axis on
back to mouse
view 3D
grid on
show big spots
axis off
grid off
show big spots
1832
1834

NIH / GLM / Tjj

GLM:tjj 09/05/02 E-272-99/2 Attorney Ref. No. 4239-62295

COMBINED DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled, the specification of which

☒ was filed on March 15, 2002 as United States Application No. 10/088,269.

☒ was described and claimed in PCT International Application No. PCT/US00/25465, filed on September 15, 2000, and amended by the Preliminary Amendment filed on March 15, 2002.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56. If this is a continuation-in-part application filed under the conditions specified in 35 U.S.C. § 120 which discloses and claims subject matter in addition to that disclosed in the prior copending application, I further acknowledge the duty to disclose material information as defined in 37 CFR § 1.56 which occurred between the filing date of the prior application and the national or PCT international filing date of the continuation-in-part application.

I hereby claim foreign priority benefits under Title 35, United States Code, § 119(a)-(d) of any foreign application(s) for patent or inventor's certificate or of an PCT International application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT International application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) on which priority is claimed:

Prior Foreign Application(s)	Country	Filing Date	Priority Claimed
			☐ Yes ☐ No

I hereby claim the benefit under Title 35, United States Code, § 119(e) of any United States provisional application(s) listed below:

Application No.	Filing Date
60/154,601	September 17, 1999

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or § 365(c) of any PCT International application(s) designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, § 1.56(a) which occurred between the filing date of the prior application and the national or PCT International filing date of this application:

Application No.	Filing Date	Status: patented, pending, abandoned

GLM:ljl 09/05/02 E-272-99/2 Attorney Ref. No. 4239-62295

I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application, to file a corresponding international application, and to transact all business in the Patent and Trademark Office connected therewith:

Name	Reg. No.	Name	Reg. No.
James C. Haight	25,588	Jack Spiegel	34,477
Susan S. Rucker	35,762	David R. Sadowski	32,808
John Peter Kim	38,514	Steven Ferguson	38,448
Stephen L. Finley	36,357	Norbert J. Pontzer	40,777
Richard R. Rodriguez	45,980	Marlene Shinn	46,005

JA

with an Associate Power of Attorney to the following:

Name	Reg. No.	Name	Reg. No.
John W. Stuart	24,540	Michael D. Jones	41,879
Richard J. Polley	28,107	Tanya M. Harding	42,630
William D. Noonan	30,878	Susan Alpert Siegel	~~43,121~~
Donald L. Stephens Jr.	34,022	Gregory L. Maurer	~~43,781~~
Wayne W. Rupert	34,420	Gillian L. Bunker	47,461
Stacey C. Slater	36,011	Anne Carlson	47,472
Stephen A. Wight	37,759	Sheree L. Rybak	47,913
Robert F. Scotti	39,830	Devon J. Zastrow	50,206
Lisa M. Caldwell	41,653		

all of the law firm of Klarquist Sparkman, LLP.

Address all telephone calls to Gregory L. Maurer, telephone number 503/226-7391 and facsimile number 503/228-9446.

Address all correspondence to:

KLARQUIST SPARKMAN, LLP
One World Trade Center, Suite 1600
121 SW Salmon Street
Portland, OR 97204-2988

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

GLM:tjj 07/25/02 E-272-99/2 Attorney Ref. No. 4239-62295

Full Name of First (or sole) Inventor: Olli P. Kallioniemi

Inventor's Signature ____________________ 8/11/02

Date

Residence: ~~Rockville, MD 20852~~ Gaithersburg, MD 20878 /OK MD

Citizenship: Finland

Post Office Address: ~~1083 Grand Oak Way, Rockville, MD 20852~~ 9601 Marathon Ter #402, Gaithersburg, MD 20878 /OK

Full Name of Second Joint Inventor: Juha Kononen

Inventor's Signature ____________________ ____________

Date

Residence: Arlesheim, Switzerland

Citizenship: Finland

Post Office Address: In Der Schappe 21, CH-4144 Arlesheim, Switzerland

Full Name of Third Joint Inventor: Lukas Bubendorf

Inventor's Signature ____________________ ____________

Date

Residence: Therwil, Switzerland

Citizenship: Switzerland

Post Office Address: Drosselstrasse 31, CH-4106 Therwil, Switzerland

Full Name of Fourth Joint Inventor: Edward R. Dougherty

Inventor's Signature ____________________ ____________

Date

Residence: College Station, TX 77845

Citizenship: United States of America

Post Office Address: 3101 Westchester Avenue, College Station, TX 77845

GLM:tj 09/05/02 E-272-99/2

Attorney Ref. No. 4239-62295

Full Name of First (or sole) Inventor: Olli P. Kallioniemi

Inventor's Signature ______________________ Date

Residence: Rockville, MD 20852

Citizenship: Finland

Post Office Address: 1083 Grand Oak Way, Rockville, MD 20852

Full Name of Second Joint Inventor: Juha Kononen

200

Inventor's Signature ______________________ 09/09/02 Date

Residence: Arlesheim, Switzerland CHX

Citizenship: Finland

Post Office Address: In Der Schappe 21, CH-4144 Arlesheim, Switzerland

Full Name of Third Joint Inventor: Lukas Bubendorf

Inventor's Signature ______________________ Date

Residence: Therwil, Switzerland

Citizenship: Switzerland

Post Office Address: Drosselstrasse 31, CH-4106 Therwil, Switzerland

Full Name of Fourth Joint Inventor: Edward R. Dougherty

Inventor's Signature ______________________ Date

Residence: College Station, TX 77845

Citizenship: United States of America

Post Office Address: 3101 Westchester Avenue, College Station, TX 77845

GLM:tjj 07/25/02 E-272-99/2 Attorney Ref. No. 4239-62295

Full Name of First (or sole) Inventor: Olli P. Kallioniemi

Inventor's Signature ______________________ Date ________

Residence: Rockville, MD 20852

Citizenship: Finland

Post Office Address: 1083 Grand Oak Way, Rockville, MD 20852

Full Name of Second Joint Inventor: Juha Kononen

Inventor's Signature ______________________ Date ________

Residence: Arlesheim, Switzerland

Citizenship: Finland

Post Office Address: In Der Schappe 21, CH-4144 Arlesheim, Switzerland

300

Full Name of Third Joint Inventor: ~~Lukas Bubendorf~~

Inventor's Signature ______________________ Date 8/5/02

Residence: ~~Therwil~~, Switzerland CHX

Citizenship: Switzerland

Post Office Address: Drosselstrasse 31, CH-4106 Therwil, Switzerland

Full Name of Fourth Joint Inventor: Edward R. Dougherty

Inventor's Signature ______________________ Date ________

Residence: College Station, TX 77845

Citizenship: United States of America

Post Office Address: 3101 Westchester Avenue, College Station, TX 77845

GLM:tjj 07/25/02 E-272-99/2

Attorney Ref. No. 4239-62295

Full Name of First (or sole) Inventor: Olli P. Kallioniemi

Inventor's Signature ______________________ ______________ Date

Residence: Rockville, MD 20852

Citizenship: Finland

Post Office Address: 1083 Grand Oak Way, Rockville, MD 20852

Full Name of Second Joint Inventor: Juha Kononen

Inventor's Signature ______________________ ______________ Date

Residence: Arlesheim, Switzerland

Citizenship: Finland

Post Office Address: In Der Schappe 21, CH-4144 Arlesheim, Switzerland

Full Name of Third Joint Inventor: Lukas Bubendorf

Inventor's Signature ______________________ ______________ Date

Residence: Therwil, Switzerland

Citizenship: Switzerland

Post Office Address: Drosselstrasse 31, CH-4106 Therwil, Switzerland

400

Full Name of Fourth Joint Inventor: Edward R. Dougherty

Inventor's Signature Ew K Ny 8/13/2002 Date

Residence: College Station, TX 77845 TX

Citizenship: United States of America

Post Office Address: 3101 Westchester Avenue, College Station, TX 77845

GLM:tjj 07/25/02 E-272-99/2

Attorney Ref. No. 4239-62295

500

Full Name of Fifth Joint Inventor: Artyom M. Grigoryan

Inventor's Signature

08/15/2002
Date

Residence: San Antonio, TX 78249

TX

Citizenship: Armenia

Post Office Address: 11246 Cedar Park, San Antonio, TX 78249

United States Patent & Trademark Office

Office of Initial Patent Examination -- Scanning Division

Application deficiencies found during scanning:

☐ Page(s)__________ of ______________________________ were not present for scanning.
(Document title)

☐ Page(s)__________ of ______________________________ were not present for scanning.
(Document title)

☑ ***Scanned copy is best available.*** Page 2 and 3 has black line
some Drawings are dark